CHARACTERIZATION OF RETROVIRAL RNA STABILITY ELEMENTS

by
Bao Lin Quek

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

June 2015

ABSTRACT

Rous Sarcoma Virus (RSV) is a simple retrovirus that relies on host cell machinery for gene expression after integration of its reverse transcribed RNA genome. Three viral RNAs are transcribed from the proviral genome, one of which is the full-length, unspliced transcript that serves as the RNA genome for progeny virions as well as the mRNA for Gag and Pol proteins. The transcript contains features uncommon in host cell mRNAs, including a long 3′ untranslated region (UTR), making it a putative target of the nonsense-mediated mRNA decay (NMD) pathway. However, this viral mRNA is protected against NMD by a small RNA element known as the RSV RNA stability element (RSE). Truncation experiments performed on the RSV RSE to identify its core element yielded partial phenotypes, suggesting the presence of redundant sub-elements across the region.

We conducted a random mutagenesis screen to determine the consensus sequences crucial for this insulator effect. Polypyrimidine tract-binding protein (PTBP1) was identified as a potential candidate that may interact with the RSV RSE to confer RNA stability. This protein was also independently identified by our collaborators in an *in vivo* RNA pull-down and mass spectrometry screen. Mutation of PTBP1 binding sites in the RSV RSE results in destabilization of the full-length transcript, and this destabilization can be rescued by re-introducing exogenous PTBP1 binding sites. However, PTBP1 does not fully account for the effects of the RSV RSE. RNA structure, as well as other proteins that may cooperatively bind the RSE or to PTBP1, may also be important for RNA stability.

As the NMD-prone features of RSV RNA are common to other retroviruses, we hypothesize that there may be similar RSE-like RNA elements present in other retroviruses as well. Regions of the 3′ UTR in Human Immunodeficiency Virus (HIV) and Moloney-Murine Leukemia Virus (MMLV) are capable of stabilizing full-length RSV transcripts in chimeric RSV constructs, suggesting that RSE-like elements exist in both these retroviruses. We have confirmed the presence of the MMLV RSE upon studying it in its natural context. However, it is unclear if HIV does have an RSE, or if it is completely immune to NMD.

Advisor/Primary reader: Karen Beemon

Secondary reader: Sarah Woodson

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor, Dr Karen Beemon. She is a great resource and wonderful mentor who has always encouraged and challenged me. Under her tutelage, I have developed into the independent scientist and critical thinker I am today.

I am grateful for the camaraderie in the Beemon lab, especially to Yingying Li for her technical expertise and companionship. I would also like to thank Gary Lam for his scientific advice, and for going above and beyond whenever I sought his help. Special thanks to Johanna Withers, who laid the groundwork for my project, and who mentored me when I first joined the lab. Thank you to my collaborators in the Bobby Hogg, Mike Summers and Wah Chiu labs, whose work was crucial in the success of my projects.

Thank you to my amazing classmates, in particular Vuong Tran and Kevin Lebo, for the joys and solidarity over the years. Thanks for being my scientific sounding boards, and for game nights and Marvel movies. You are the reason I made it through graduate school, but also the reason why I did not graduate sooner.

To my dear friends, especially Yorke Zhang, Hong Yuen Wong, and Bing Shao Chia. Thank you for the laughs, rants, invigorating debates, and mindless chitchat. Thanks for sticking with me and believing in me even when I didn't believe in myself.

Finally, thank you to my family for encouraging me to excel and pursue my dreams. Their never-ending love and support is the biggest reason for my success. 感谢我的家人给予我的爱护和支持，不断的鼓励我去追逐梦想，展翅而飞。尤其是妈妈，在我成功时帮我喝彩，在我苦恼时替我分忧。多亏你，我才有今天的成就。谢谢！

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1:

INTRODUCTION

Adapted from Quek BL, Beemon K. (2014) Retroviral strategy to stabilize viral RNA. Curr. Opin. Microbiol. 18, 78-82.

An overview of the retroviral life cycle

Viruses are obligate parasites with compact genomes that depend on host cells for replication. Retroviruses are single-stranded, + strand RNA viruses with a double-stranded DNA intermediate. Figure 1 [1] depicts an overview of the retroviral life cycle [2], using Human Immunodeficiency Virus (HIV), a complex retrovirus, as an example. Two copies of its genome, along with viral proteins such as protease, reverse transcriptase (RT) and integrase (IN), are packed into the viral capsid and surrounded by the envelope.

Upon infecting a new cell, the surface viral glycoproteins bind receptors on the host cell plasma membrane, triggering a membrane fusion event and the release of the viral core into the cytoplasm. There, the capsid partially disassembles, allowing dNTPs to enter the core. RT uses the viral RNAs as templates to make double stranded DNA, and forms a pre-integration complex (PIC) with host and viral proteins. For HIV and RSV [3–5], this complex gets imported into the nucleus through the nuclear pore, but for Moloney Murine Leukemia Virus (MMLV), the PIC stays in the cytoplasm and only enters the nucleus when the cell undergoes mitosis. Once in the nucleus, the double stranded proviral DNA gets integrated into the host cell genome with the help of IN.

Thereafter, the provirus essentially hijacks host cell mechanisms to make more viruses. The proviral DNA is transcribed by RNA pol II, and the viral RNAs are spliced and processed just like other host cell transcripts. Ribosomes translate RNAs to synthesize various viral proteins, which are then transported to their appropriate locations. Finally, Gag proteins and viral RNA genomes are assembled near the plasma

membrane, followed by viral budding, although there is evidence that RSV assembly starts in the nucleus [6, 7].

For a retrovirus to successfully complete its life cycle, it has to avoid various host cell defense mechanisms, such as the RNA surveillance system. This surveillance system consists of three major components: non-stop decay [8–10], which detects transcripts lacking a stop codon that are translated through the polyA tail; no-go decay [11–13], whereby translating ribosomes are unable to read through a strong secondary structure in the RNA, causing the ribosome to stall and trigger decay; and nonsense-mediated mRNA decay (NMD) [14–16], which presents the biggest threat for retroviruses.

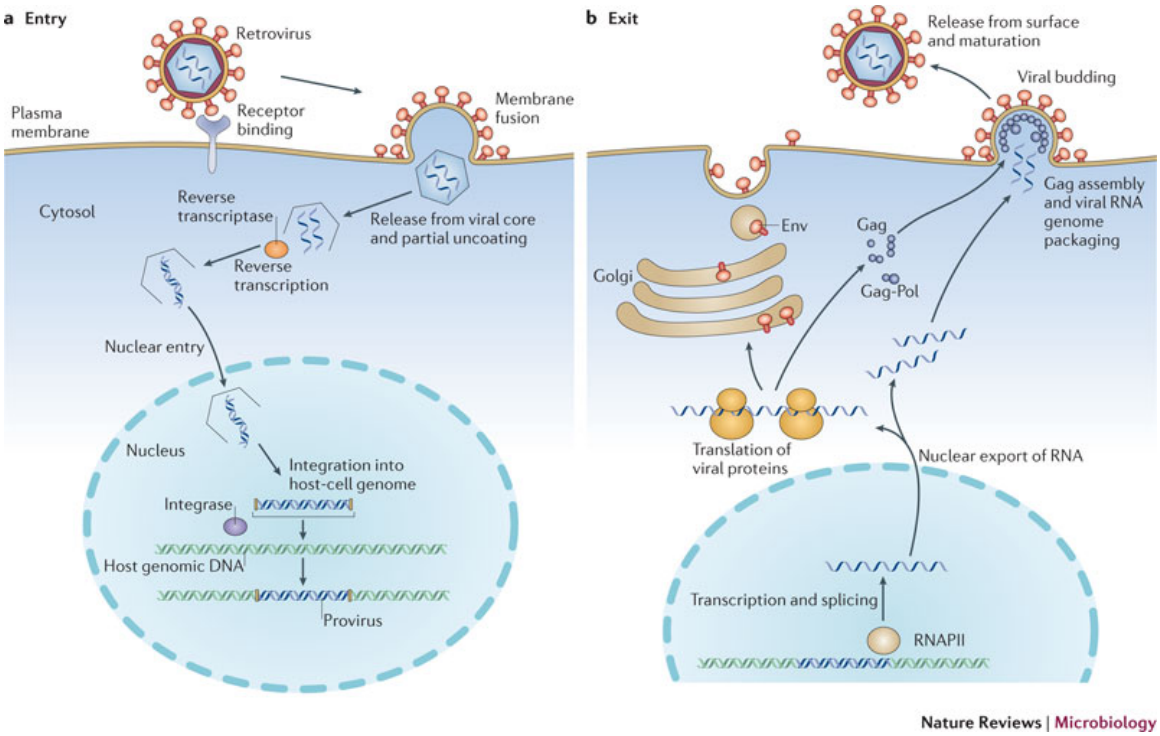


Figure 1: Major events in the retroviral life cycle. **(a)** Viral entry into cells begin with the binding to a specific receptor on the cell surface. Membrane fusion occurs either at the plasma membrane or from endosomes (not shown). The viral core is released into the host cytoplasm, whereby the capsid partially uncoats. dNTPs enter the viral core, initiating reverse transcription of the viral genome. The pre-integration complex then transits through the cytoplasm, enters the nucleus, and the proviral DNA integrates into the cellular DNA. **(b)** Viral exit begins with proviral DNA transcription by RNA polymerase II (RNAPII). The viral RNA is incompletely spliced and exported to the cytoplasm, where translation of viral proteins occurs. Gag assembly and viral RNA packaging occurs, followed by viral budding through the cell membrane. The virus is then release from the cell surface and undergoes maturation by protease.

Features of retroviral RNAs predispose them to degradation by NMD

Retroviruses have compact genomes of less than 10 kilobases (kb), yet code for 8 or more viral proteins. In contrast, most single cellular genes are much larger than this, due to numerous large introns. Retroviruses use a combination of gene expression mechanisms, including alternative splicing, frame-shifting and polyprotein cleavage to generate multiple gene products from a single primary RNA transcript [17]. The major unspliced retroviral mRNAs have features not commonly found in cellular mRNAs, including long 3′ untranslated regions (3′ UTRs), retained introns, and upstream open reading frames (uORFs). These unusual features predispose the viral RNA towards degradation by NMD [18–21]. Nevertheless, the unspliced RNA of RSV, a well-studied avian retrovirus, is very stable [22], suggesting the virus has evolved a means to overcome degradation by NMD.

An overview of NMD

NMD was originally proposed to be responsible for degrading aberrant transcripts containing premature termination codons (PTCs), thus protecting cells from the deleterious effects of C-terminal truncated proteins [23]. Now, however, it also seems to provide a mechanism for regulating levels of expression of alternatively spliced mRNAs[21], and serves a regulatory role across the transcriptome [24].

NMD is a translation termination event; decay is inhibited when translation is halted. The key component of this pathway is the ATP-dependant RNA helicase Up-

frameshift 1 protein (Upf1) [25–28]. Smg1, which phosphorylates Upf1, is also required. Other factors, such as Upf2 and Upf3 may also be involved in triggering NMD, and decay is mediated via Smg5, Smg6, and Smg7, as well as other decay factors such as Xrn1 [20, 21]. However, there are several mechanistically distinct branches of NMD, and the factors required and involved for the different branches vary [29].

Are downstream EJCs necessary for NMD?

Canonical NMD in mammalian cells posits that the process is mediated by the interaction between the terminating ribosome at a PTC and a downstream exon-junction complex[26] (EJC), the multi-protein complex deposited 20-24 nts upstream of the exon-exon junction during RNA splicing [30, 31] (see Fig. 2b, compared to 2a). However, examples of EJC-independent NMD exist in mammals and are the rule in lower organisms such as yeast [32]. Some, but not all, mRNAs bearing features such as uORFs and long 3′ UTRs are NMD targets [18–21], both in mammals and in yeast. Unspliced (and presumably lacking EJCs) avian retroviral mRNAs are also targets of NMD [33]. NMD has been shown to play regulatory roles in humans, Drosophila melanogaster, Caenorhabditis elegans, and Saccromyces cerevisiae [21], and can affect more than a quarter of all expressed mammalian genes [34, 35].

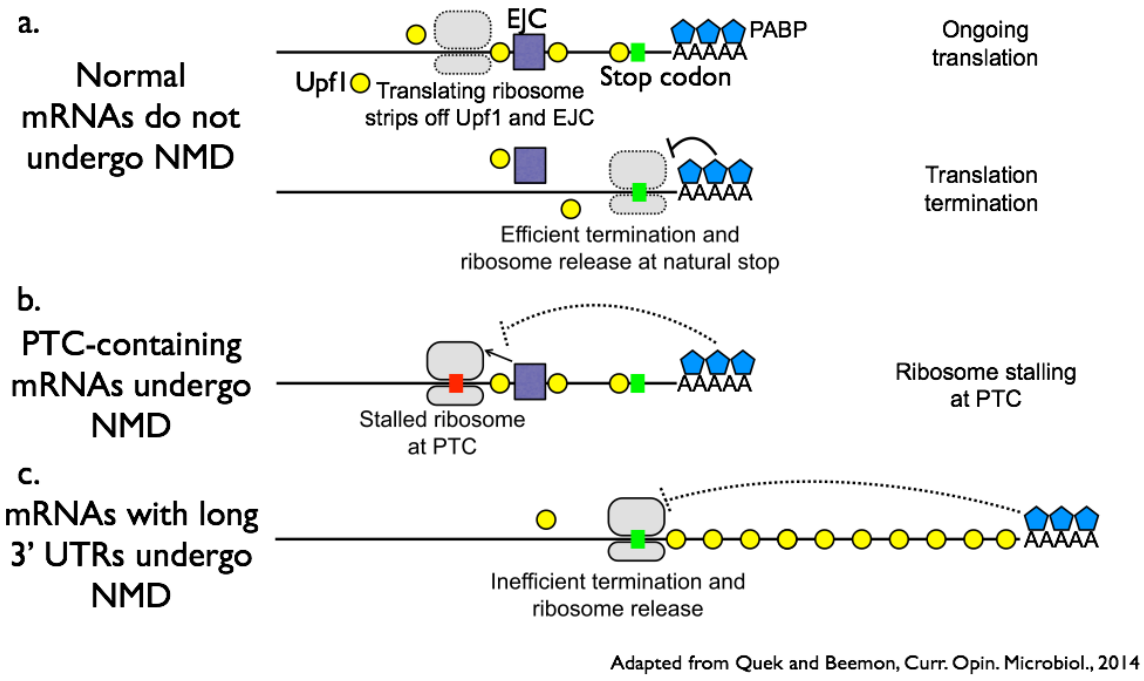


Figure 2: NMD susceptibility of various transcripts. (a) A normal mRNA is bound by EJCs in the coding region and by Upf1 across the transcript. The translating ribosome strips off Upf1 and EJCs from the coding sequence. At the natural stop codon, translation termination is efficient. Coupled with interactions with the proximal PABPs, ribosome disassembly is rapid and NMD factors are not recruited. The mRNA is stable. **(b)** At a premature termination codon (PTC), translation termination is inefficient. The distance to the poly(A) tail prevents PABP intervention, and the stalling ribosome interacts with downstream EJCs and Upf1 to signal NMD. The transcript is degraded. **(c)** For mRNAs with long 3′ UTRs, the terminating ribosome is far from the poly(A) tail and PABP. Additionally, increased Upf1 binding in the 3′ UTR increases chances of interaction with the ribosome. This makes the mRNA susceptible to NMD.

Metze et al [29] compared “EJC-enhanced” and “EJC-independent” NMD in mammalian cells, and showed that they represented two partially redundant degradation pathways, each involving different NMD factors. EJC-independent NMD shows a stronger dependence on Upf2 and Upf3. NMD targets are degraded either by Smg6 endonuclease or Smg5/Smg7, which interact with each other and appear to have exonuclease activity. However the inhibition of a decay pathway leads to the usage of the other pathway. Nonetheless, all PTC-containing mRNAs are degraded in a Smg1- and Upf1- dependent manner, regardless of the presence or absence of a downstream EJC. The EJC could therefore serve as a non-essential, indirect NMD enhancer by increasing translation efficiency, a role that has previously been demonstrated [36], along with its role in mRNA export.

Saulière et al [37] and Singh et al [38] recently monitored EJC binding across the transcriptome in HeLa and HEK293 cell lines, respectively. These studies showed that canonical EJCs (cEJCs) are only detected at 80% of exon junctions and at varying occupancy levels, so that many individual mRNAs lack EJCs at many exon junctions. Surprisingly, almost half of all mapped EJC binding sites were found distributed across the mRNA coding sequence rather than at the exon junctions [37, 38]. Furthermore, Saulière et al observed an enrichment of the serine/arginine-rich (SR) protein consensus binding motif GAAGA near the mapped EJC binding sites, and showed that SRSF1 (ASF/SF2) and SRSF7 (AAG3/9G8) physically associate with EJCs [37]. SR proteins are RNA binding proteins that recruit splicing factors. Correspondingly, Singh et al found that multiple EJCs were often associated with many SR proteins in large complexes [38]. This interaction between EJCs and SR proteins could explain the previously observed

promotion of NMD by overexpression of the SR protein SRSF1 [39]. This blurs the distinction between the two branches of NMD [29], as transcripts with downstream introns may not have EJCs occupying them and transcripts thought to undergo EJC-independent NMD may in fact possess non-canonical EJCs (ncEJCs). Most transcripts carry both cEJCs and ncEJCs [37, 38]. However, EJCs are rarely seen in the UTRs or in intronless mRNAs [37, 38].

NMD studies have also been done on two togaviruses, the Semliki Forest virus (SFV) and Sindbis virus (SINV), which are + strand RNA viruses that replicate in the cytoplasm. SFV and SINV viral RNAs are polycistronic RNAs with long 3′ UTRs that undergo NMD upon entering the cell [40]. These RNAs do not enter the nucleus and presumably do not have EJCs; there is no evidence of EJCs loading onto RNAs in the absence of splicing, even though EJC proteins are present in the cytoplasm. This lends further support to the hypothesis that EJCs are not required for NMD.

PABP stabilizes RNA

A different model was put forward to explain how yeast cells discriminate between normal and aberrant termination codons. (See Fig. 2c, compared to 2a.) Long 3′ UTRs were shown early on to be associated with NMD in yeast, leading to the faux UTR model that suggests these UTRs promote NMD because they lack termination regulatory factors normally present on legitimate 3′ UTRs [41]. Since cytoplasmic poly(A) binding proteins (PABPC1) tethered near a terminating ribosome can prevent NMD [42], it was proposed that long 3′ UTRs induce NMD by distancing the PABPs from the terminating

ribosome. Studies with recombinant mRNAs in Drosophila also showed the importance of a short distance between a termination codon and poly(A) for stable mRNA [43]. While *Drosophila* has EJCs, they do not seem necessary for NMD in flies [43]. An artificial mammalian mRNA with a long 3′UTR could be protected from NMD by looping the 3′ UTR via base-pairing to bring the poly(A) tail close to the termination codon [44]. Similarly, tethering PABP near a PTC also stabilizes the mRNA. [44, 45]

How does UPF1 recognize target RNAs?

Despite being the key component of the NMD pathway, it is not well understood how Upf1 recognizes and binds its targets. Recent work has attempted to answer that question. Hogg and Goff showed that Upf1 directly binds to 3′ UTRs of several mRNAs (HIV-1, Smg5 and GAPDH) in a length-dependent, translation-independent manner [46]. In contrast, Kurosaki and Maquat observed that Upf1 binding to non-PTC-containing ß-globin constructs does not correlate with 3′ UTR length [47]. Furthermore, transcripts that contain PTCs or downstream EJCs demonstrate augmented Upf1 binding, and this enhanced binding depends on translation. They propose that Upf1 is loaded onto the 3′ UTRs of PTC-containing mRNAs at the terminating ribosome. [47]

Two recent studies interrogated Upf1 binding across the transcriptome in murine embryonic stem cells (mESCs) and HeLa cells via crosslinking/immunoprecipitation-sequencing (CLIP-seq) and individual-nucleotide-resolution UV cross-linking and immunoprecipitation (iCLIP) [48, 49]. Both studies agree that Upf1 preferentially binds the 3′ UTR at up to 10 times higher levels than the coding sequence. The binding is

relatively uniform across the 3′ UTR; however, both groups observed clustering in specific locations in individual transcripts. In addition, Hurt et al [49] observed a binding preference for G-rich sequences and RNA regions with secondary structure, while Zünd et al [48] noted enrichment at U-rich sequences. Notably, translation inhibition causes a dramatic increase in Upf1 binding to the coding sequence [48, 49]. Upf1 also binds long nonconding RNAs (lncRNAs). Thus, in contrast to current models proposing that Upf1 is specifically recruited to the 3′ UTR by ribosomes, these studies suggest that Upf1 is uniformly bound across transcripts and is displaced by the translocating ribosome. This implies that Upf1 binding happens prior to the selection of NMD targets. Accordingly, Zünd et al do not observe a preference for Upf1 binding to NMD targets. [48]

Nonetheless, it is still unclear how 3′ UTR length functions as a determinant of NMD susceptibility in this model, and how increased Upf1 binding promotes NMD and increases decay efficiency. One possible model is that Upf2 and Upf3 find and bind the RNA-bound Upf1 through diffusion. [21] This primes the Upf1-Upf2-Upf3 complex for signaling decay upon interaction with the terminating ribosome. Longer 3′UTR's usually bind more Upf1 [50] and have an increased probability of interacting with diffusing Upf2 and Upf3 factors.

Alternative polyadenylation and other mechanisms to modulate 3′ UTR length

Cellular mechanisms exist to modulate 3′ UTR length, which may affect mRNA localization and translational efficiency in addition to its stability. Deletions within the 3′ UTR, as well as point mutations that lead to premature cleavage and polyadenylation,

alter 3′ UTR length and sequence. [51, 52] Chromosomal translocations have also been shown to result in 3′ UTR swapping between genes. [53] In addition, alternative cleavage and polyadenylation (APA) is a recently studied post-transcriptional regulation. [53–55] Most genes contain multiple polyadenylation signals, in both the 3′ UTR and the coding sequence. Use of an APA signal in the coding sequence will generate different protein isoforms, whereas use of an APA signal in the 3′ UTR will yield mRNA isoforms with varying 3′ UTR lengths and associated regulatory elements. In humans, APA signals are used in half of all genes. [56, 57]

Retroviral mechanisms to avoid NMD

Unlike mammalian mRNAs, retroviral mRNAs cannot modulate their stability via 3′ UTR length switching because they must maintain full-length genomic RNA. The full-length, unspliced transcript serves as the RNA genome for progeny virions as well as the mRNA for structural (Gag) and enzymatic (Pol) viral proteins. [17] Consequently, some retroviruses have developed protein-based or RNA-based protective mechanisms against NMD.

Human T-lymphotrophic virus type 1 (HTLV-1) Tax protein binds Upf1, as well as INT6, a subunit of the translation initiation factor eukaryotic initiation factor 3 (eIF3) required for efficient NMD, and sequesters them to ensure viral RNA integrity and translation. [58] In addition, Rex establishes a general block to NMD to stabilize both viral and host cell transcripts. [59] Of the cellular or viral mRNAs described to evade

NMD, the most extensively characterized is a cis-acting RNA sequence found in the Rous sarcoma virus (RSV), designated the RNA stability element (RSE). [22, 60, 61]

RSV is a simple avian retrovirus with three viral RNAs [62] generated from the primary transcript by alternative splicing. The full-length RSV mRNA contains several NMD-triggering features, including upstream ORFs and a long 3′ UTR. However, the RNA is stable in chicken cells and has a long half-life of about 20 hours [22], suggesting it has evolved a mechanism to avoid degradation by NMD.

Studies of mouse HSP70 and human histone H4 unspliced mRNAs found that they were resistant to NMD [63], leading to the proposal that all unspliced mammalian mRNAs are resistant. On the other hand, most S. cerevisiae mRNAs are not spliced, yet they are subject to NMD [20]. Additionally, RSV unspliced RNA with frameshift mutations or PTCs in the gag gene is rapidly degraded in a Upf1- and translation-dependent manner, showing that the transcript is recognized as an NMD substrate, albeit in chicken cells. [22, 64]

Neither the EJC model from mammalian cells or the long 3′ UTR model from yeast and drosophila seemed adequate to explain the mechanism of NMD of RSV mRNA. Since the mRNA is unspliced, it was not thought to have EJCs. In addition, the normal gag termination codon is 7 kb away from the poly(A) tail, generating a very long 3′ UTR but a stable RNA. In contrast, insertion of a PTC 100 nts or more upstream of the normal termination codon results in degraded mRNA. [22] Thus, the distance from poly(A) did not seem to be a key determinant here.

The finding that PTCs within 100 nts of the normal *gag* termination codon showed partial RNA stability [22] suggested that the sequences downstream of the termination codon could be significant in creating a proper context for translation termination. Deletion experiments led to the identification of a 400 nt cis-acting RNA element in RSV located immediately downstream of *gag*, termed the RSV RNA stability element (RSE). [65] The RSE, which contains redundant sub-elements and a minimal 155 nt core, stabilizes the full-length mRNA and protects it from NMD. [65, 66] Deletion or inversion of this sequence destabilizes the RNA; stability can be rescued by expression of a trans-dominant Upf1 mutant to inactivate the NMD pathway. [65] Insertion of the RSE after a PTC in *gag* also rescues the RNA from decay. [65]

Alignment of 20 different avian retrovirus strains shows that the sequence and secondary structure of the RSE is well conserved. [67] This suggests that the RSE may be an RNA element that certain viruses have evolved to evade host cell RNA surveillance and to stabilize their genome. In addition, the mechanism used by the RSE could perhaps already subsist as another facet among the repertoire of RNA regulation methods in a cell. RNAs with long 3′ UTRs that escape NMD occur naturally in the cell, for instance Cript1 and Tram1. [45] It is possible that they may also have stability elements downstream of their termination codons. In support of this hypothesis, a recent study showed that for a subset of human mRNAs with long 3′ UTRs, there is a cis-acting element within 200 nts downstream of the natural termination codon that is important for RNA stability. [68]

Research objective

The RSV RSE confers a novel mode of protection from NMD. One of the main goals of my project was to determine its exact mechanism of protection. To achieve this goal, I adopted a mutagenesis approach to identify the key sequences in the RSE that are crucial for its role in conferring RNA stability. From there, I investigated factors that bind the RSE and are involved in RSE-mediated RNA stability. Another goal of my project was to determine if this mode of protection against NMD is conserved across other retroviruses. For this, I investigated the potential of 3′ UTRs of Moloney Murine Leukemia Virus (MMLV) and Human Immunodeficiency Virus (HIV-1) in stabilizing chimeric RSV. Thereafter, I deleted their putative RSEs in their natural context to see if they lose protection of their viral genomes from NMD.

CHAPTER 2:

CHARACTERIZING THE ROUS SARCOMA VIRUS

RNA STABILITY ELEMENT (RSV RSE)

Identifying sequences in the RSV RSE that are crucial for protection against NMD

Figure 3(a) shows the secondary structure of the RSV RSE RNA as predicted by selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) chemistry and RNase digestion. [67] Previous work on the RSV RSE has shown that there are redundant sub-elements across its 400 nts, especially in the 5′ and 3′ regions. [65] Weil showed that the 5′-most 250 nt fragment (nts 2486–2735) and the 3′-most 280 nt fragment (nts 2606–2885) of the RSE can partially stabilize full-length RSV RNA, although the overlapping 130 nt region (2606–2735) is necessary but insufficient for any stabilization effects. Withers expanded on that work to show that the minimum central region necessary for protection against NMD spans 155 nts, and designates the 2578–2732 nt region as the minRSE. [66] To identify all of these sub-elements across the RSE and to parse apart their individual contributions to full-length RSV RNA stability against NMD, I decided to carry out a random mutagenesis screen to introduce point mutations into the element. The goal of the screen is to obtain loss-of-function mutants, and I hypothesized that the mutations in the mutant library will cluster at regions important for stability.

The screen was performed with two different constructs. The first construct (see Figure 3b) consists of the C-terminal end of the RNA fragment (C-frag), nts 2578–2885, which encompasses the entire element except for the ribosomal frameshift pseudoknot. Previous truncation data has shown that the pseudoknot plays no role in RSV RNA stability [66]; hence I excluded the region from our mutagenesis to leave the frameshifting machinery intact while still covering all other redundant regions. The second construct used (see Figure 3c) was the 155 nt central core of the RSE, spanning

nts 2578–2732 and denoted minRSE, which is the smallest fragment of the RSE that is able to recapitulate wildtype (WT) protection effects against NMD when inserted after the natural stop codon (NTC). [66] The minRSE minimized redundancy, and sped up my search for important consensus sequences and potential interactors with the RSE. Point mutations were introduced into the C-fragment RSE and minRSE by mutagenic PCR, and the mutated fragments were respectively cloned back into full-length RSV RNA via the EagI and SpeI restriction sites previously engineered into the RSV WT E/S 2.0 vector (Figure 3b) and the RSV WT E/S vector [66] (Figure 3c). Figure 3b and 3d depict the final constructs for C-frag mutagenesis and minRSE mutagenesis.

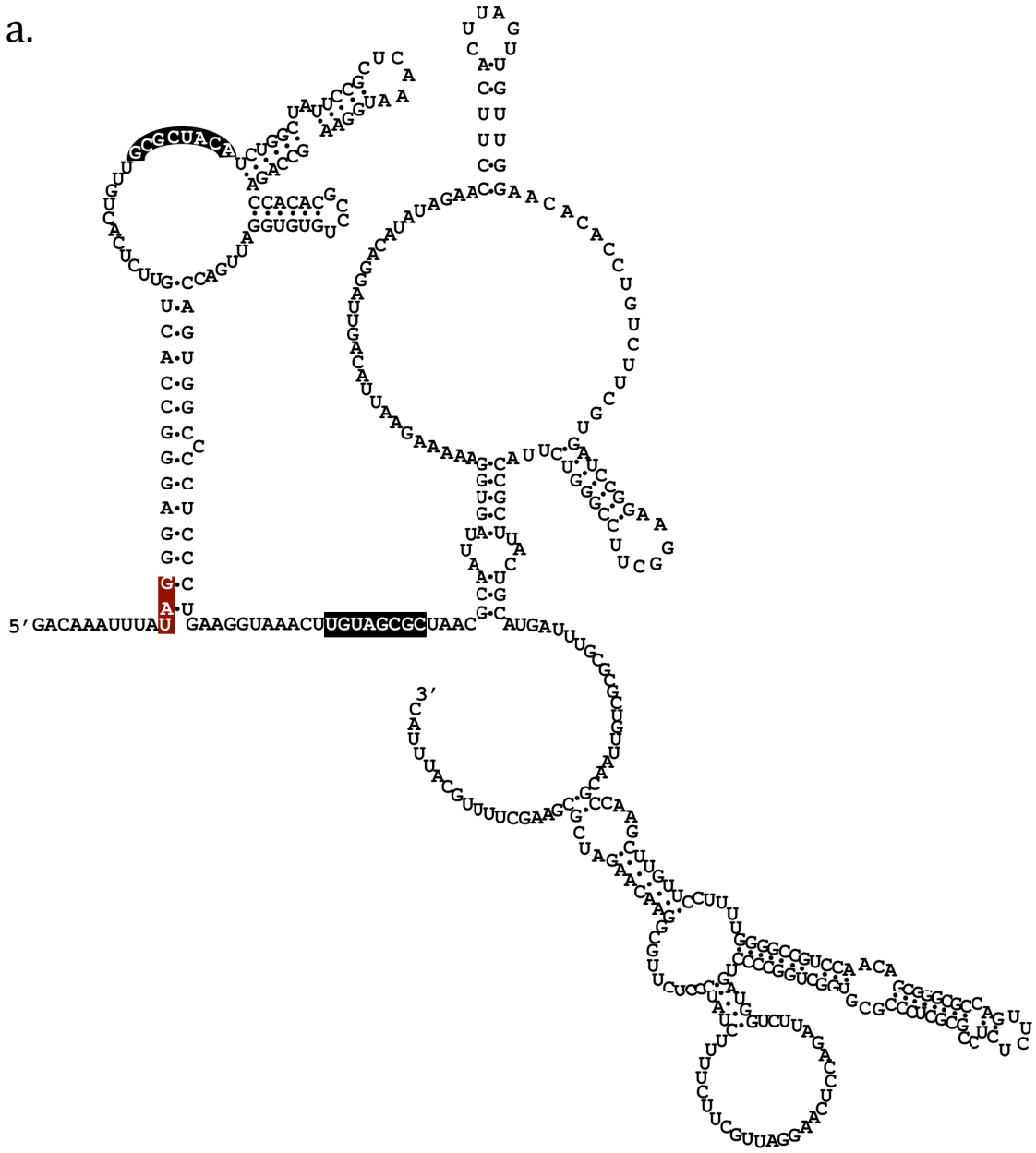


Figure 3(a): WT RSV RSE. Secondary structure of WT RSV RSE as determined by selective 2′-hydroxyl acylation analyzed by primer extension (SHAPE) chemistry and RNase digestion. *Gag* stop codon marked in red box (nts 2483–2485), frameshift pseudoknot base pairing marked in black boxes (nts 2509–2516 and 2595–2602).

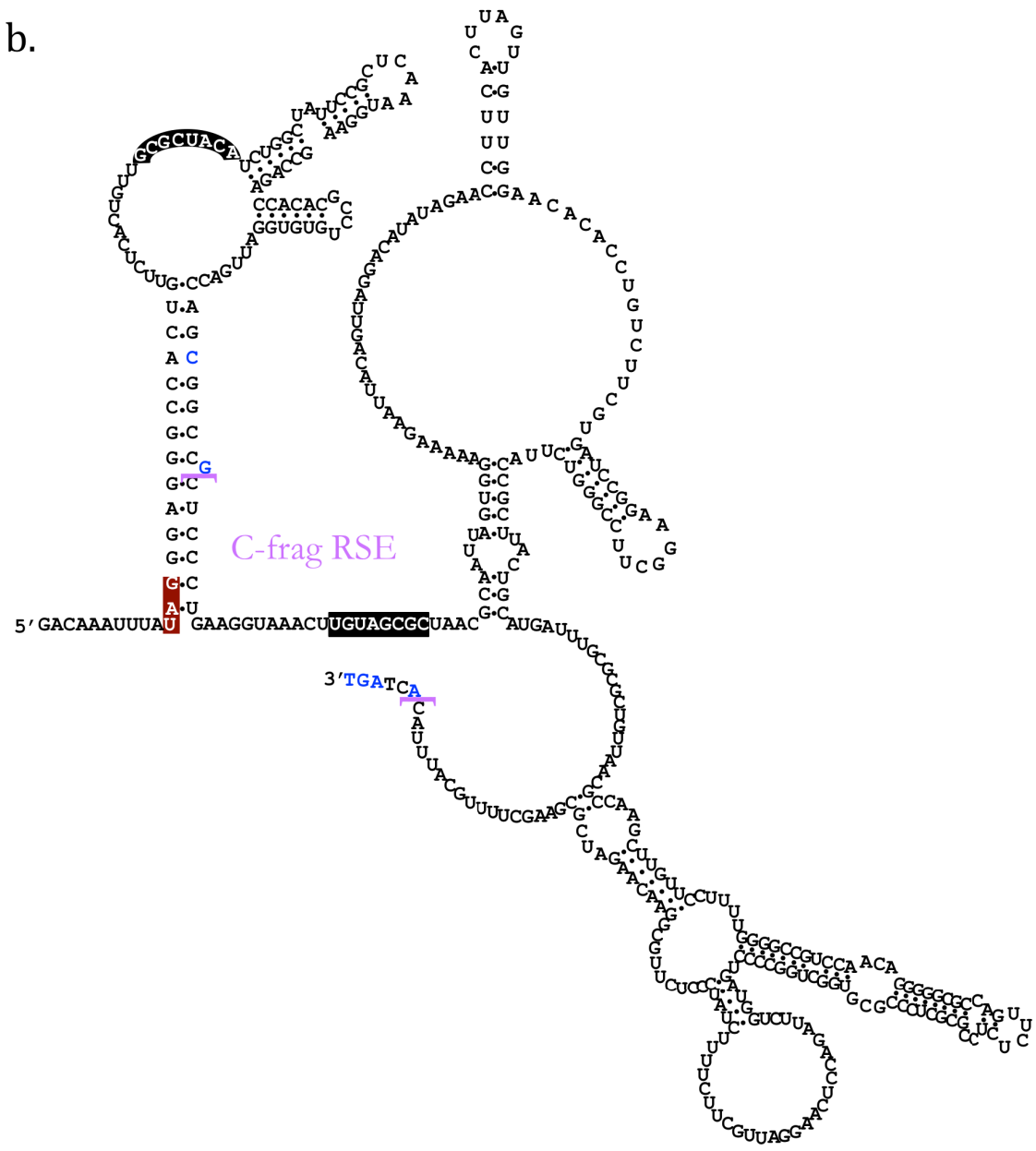


Figure 3(b): C-frag RSE mutagenesis in RSV WT E/S 2.0 vector. Purple brackets mark boundaries of C-frag RSE. RSV WT E/S 2.0 vector was used for mutagenesis. Blue nucleotides denote point mutations made to generate the 5′ CGGCCG EagI restriction site and the 3′ ACTAGT SpeI restriction site.

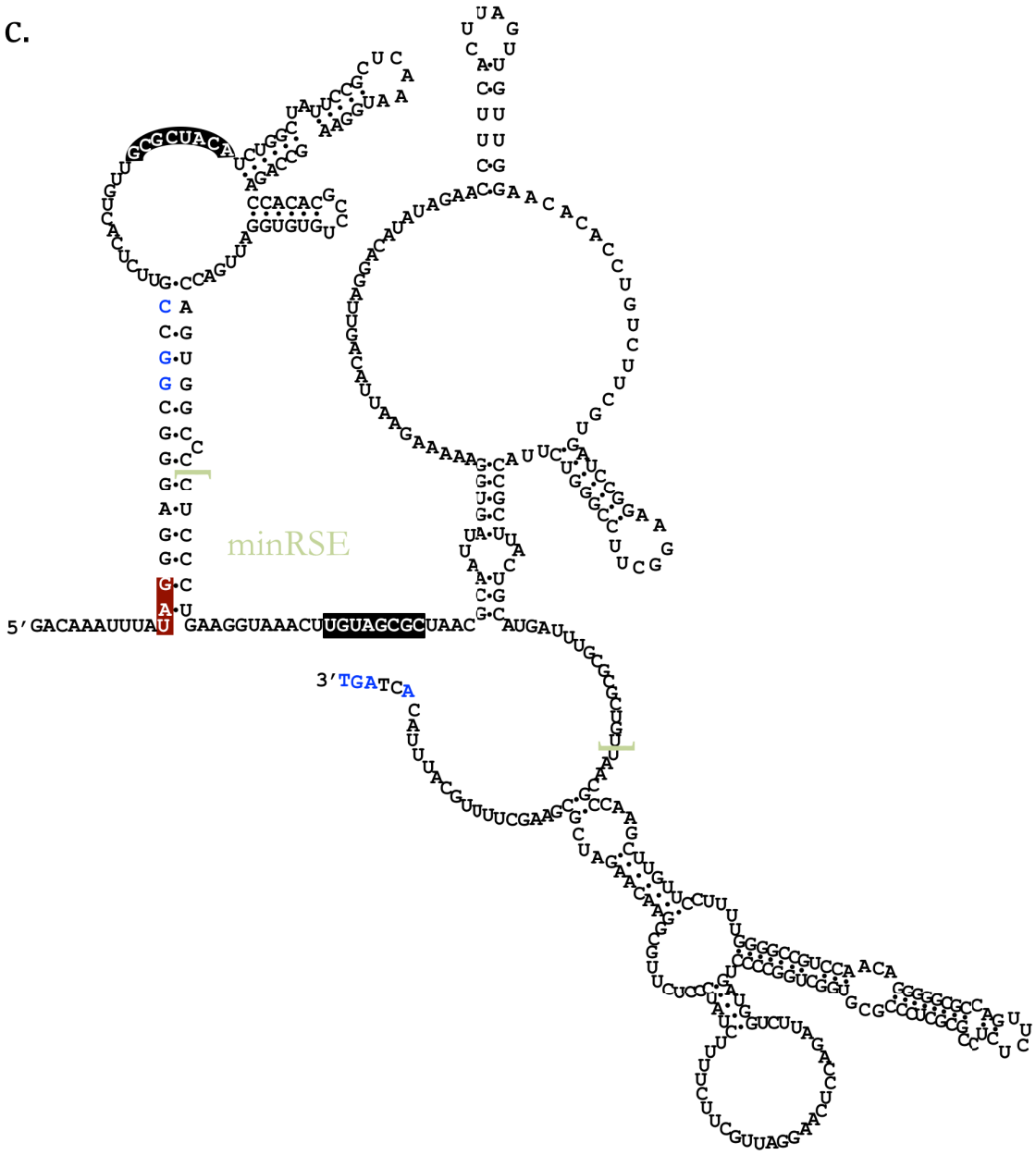


Figure 3(c): RSV WT E/S vector. Green brackets mark boundaries of minRSE. Blue nucleotides denote point mutations made to generate the 5′ CGGCCG EagI restriction site and the 3′ ACTAGT SpeI restriction site.

d.

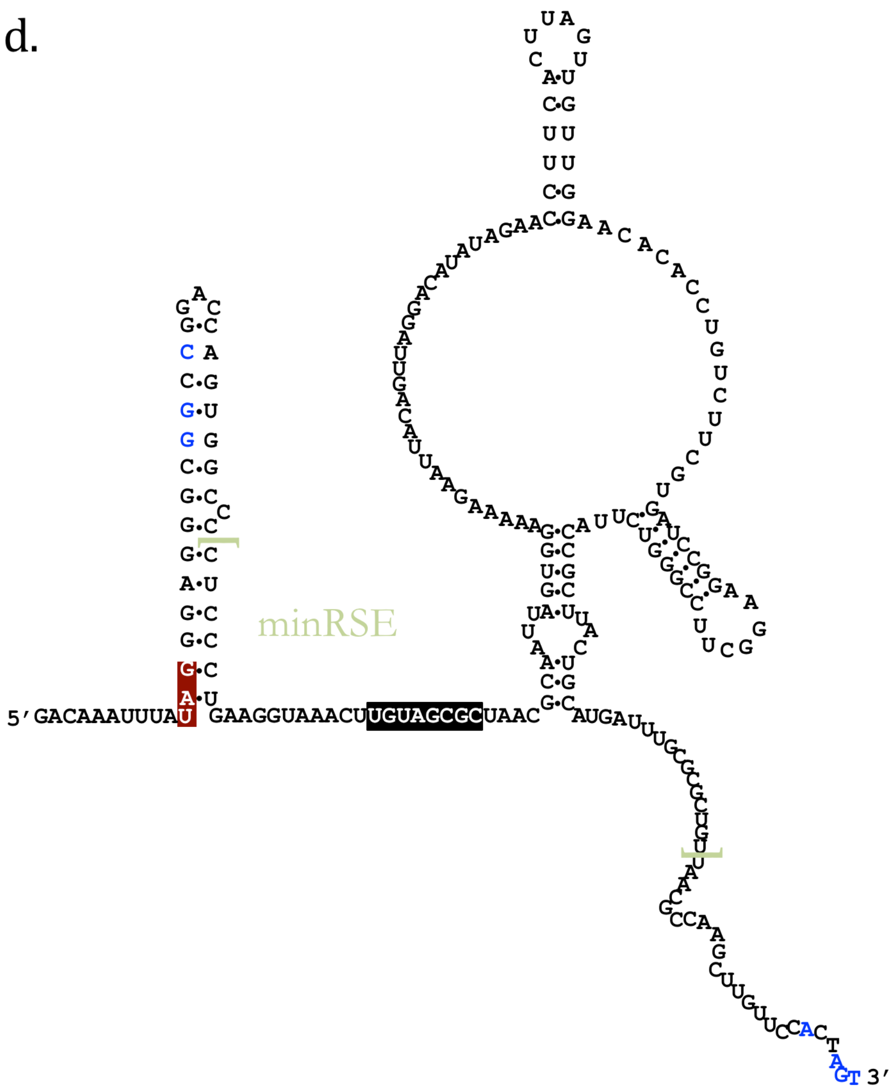


Figure 3(d): minRSE mutagenesis in RSV WT E/S vector. Green brackets mark boundaries of minRSE. Blue nucleotides denote point mutations made to generate the 5′ CGGCCG EagI restriction site and the 3′ ACTAGT SpeI restriction site. The additional WT RSV RSE nucleotides between the restriction sites and outside the boundaries of the minRSE serve as primer binding sites that are largely unchanged during mutagenic PCR. This ensures that all nucleotides of the minRSE can and will be mutated.

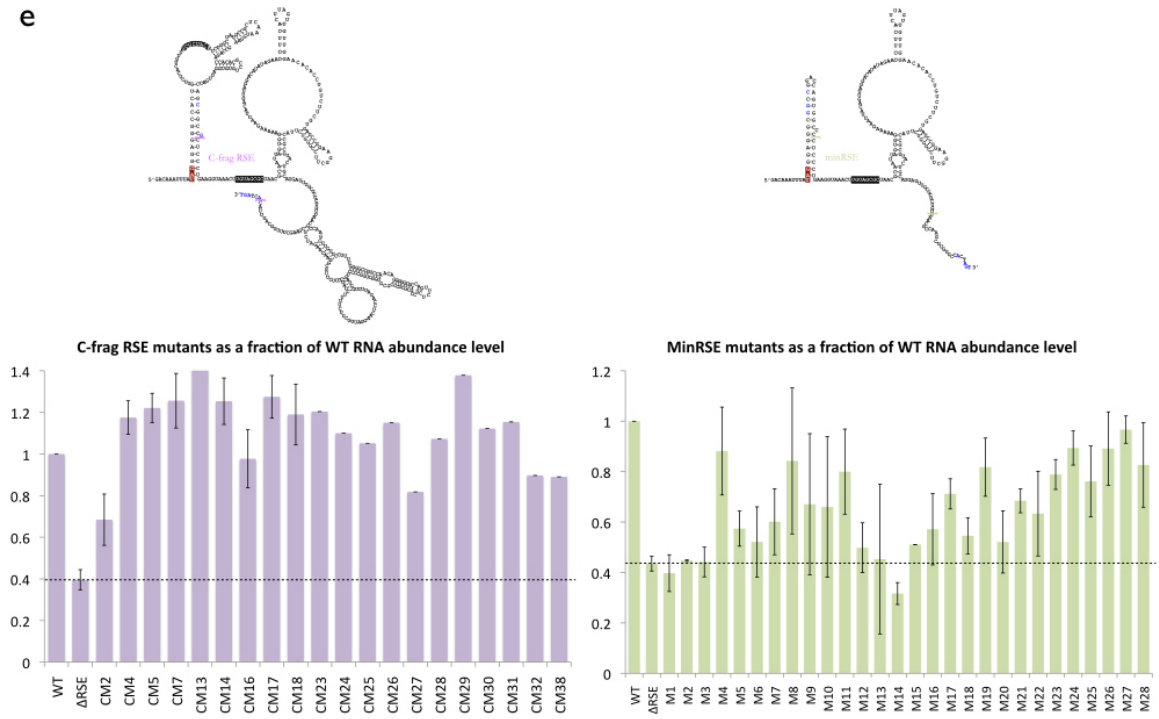


Figure 3(e): Random mutagenesis of RSV RSE to identify consensus sequences and screen for interaction partners. **Left:** RNase protection assays (RPAs) on C-frag RSE as marked in Fig 3b. C-frag mutants display a range of stabilization effect, and are generally more stable than minRSE mutants as expected. **Right:** RPAs on minRSE as marked in Fig 3d. minRSE mutants have different and multiple point mutations, and show a range of stabilization effect. Standard error represented on graph, $n \geq 2$.

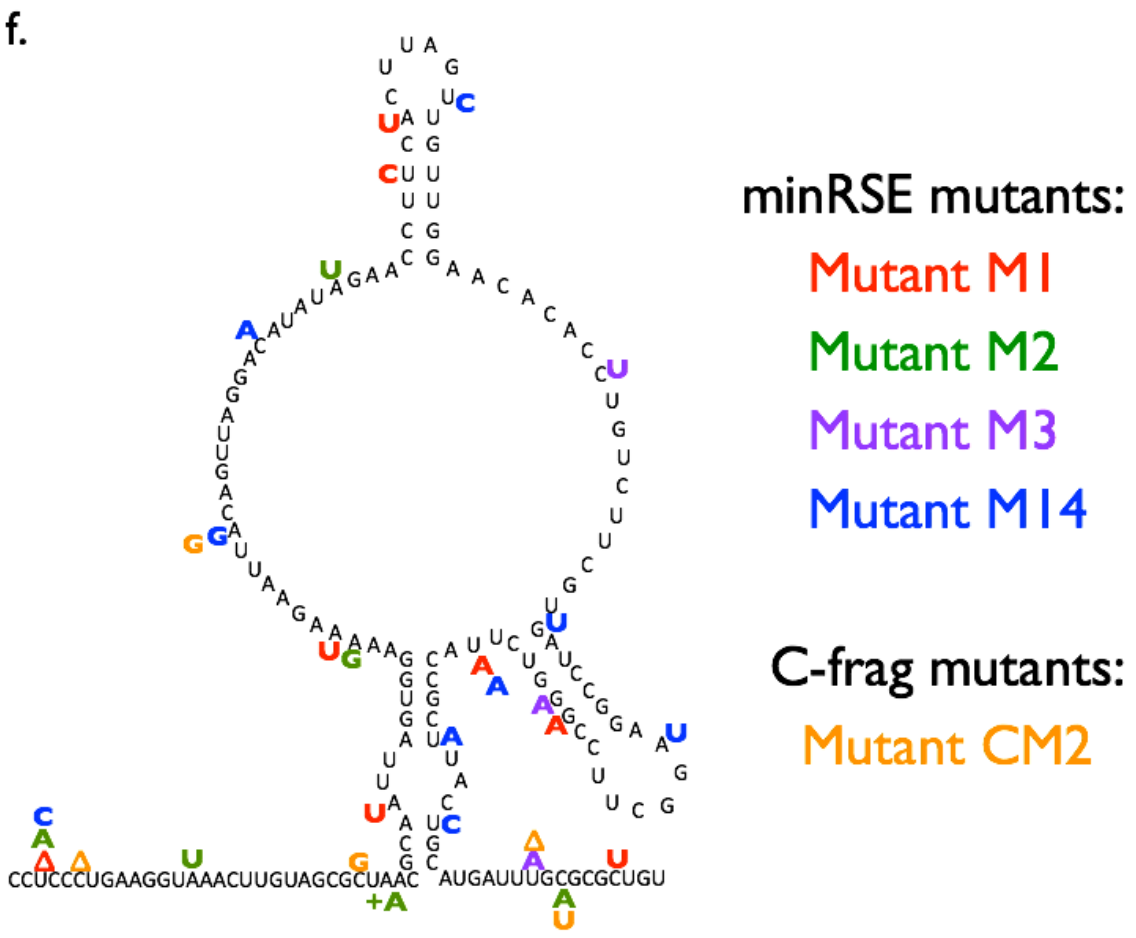


Figure 3(f): Loss of function mutations in RSV RSE. minRSE mutants that showed RNA abundance levels at or lower than ΔRSE as per Fig 3e were mapped. C-frag mutant CM2, which showed the lowest RNA abundance levels in the C-frag mutagenesis, was also mapped to this minRSE structure; CM2 only had one other mutation in addition to the ones shown here, located in the 3′ GC-rich stem loop.

Figure 3e shows the results of the screen. Steady-state viral RNA abundance levels were measured and used as a proxy for RNA stability in RNase protection assays (RPAs). As compared to the WT, most C-frag mutants exhibited full or partial functionality, which is to be expected since it would be rare to concurrently mutate all important regions in a single fragment. However, CM2 does show a decrease in RNA stability. Most minRSE mutants exhibit a much more severe phenotype, including mutants M1, M2, M3, and M14, which have complete loss of function.

I then mapped the mutations from LM1 and the minRSE loss-of-function mutants back onto the sequence and structure of the minRSE (see Figure 3f). There were regions where mutations seemed to cluster, with 3 or more of the 5 mutants carrying mutations in that region, as well as a few specific nucleotides that were frequently mutated. I was particularly interested in the 5′ and 3′ ends of the minRSE, as the deletion of these specific regions caused large destabilization effects as demonstrated in Wither's truncation experiments. [66]

PTBP1 is a candidate protein that may interact with RSV RSE

Focusing on these clusters, I performed a computational analysis, browsing the RNA Binding Protein Database (RBPDB) in search of potential protein interactors that may bind the RSE. The RBPDB is a curated database of RNA-protein interactions in human, mice, flies, and worms, as reported in primary literature. [69] Using a 6 nt motif centered on the each of the mutation clusters, I obtained a list of candidate proteins, which are listed in Table 1. This search yielded a surprisingly short list, with very little

overlap between the clusters. Polypyrimidine tract binding protein 1 (PTBP1) was the most interesting candidate as it bound to 12 of the 14 motifs tested. Other interesting candidates include serine/argenine-rich splicing factor 1 (SFRS1), which bound to 5 of the 14 motifs tested, as well as human antigen R (ELAV1). ELAV1 bound to 5 of the 14 motifs tested, and was also previously identified by Withers in an MS2 pulldown/mass spectrometry screen (unpublished).

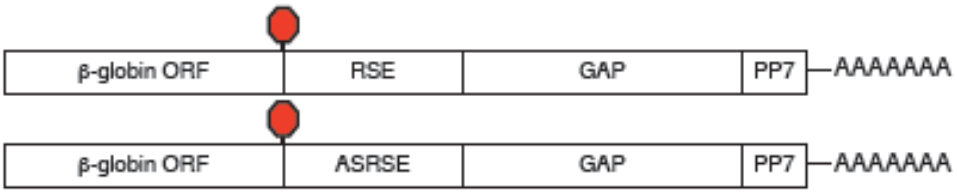
CUCCCU	GCUAAC	ACGCAA	AAAAAG	UUACAG	CAUAUA	CUUCAC
PTBP1	CSDE1	PTBP1	PTBP1	PTBP1	PUM1	PTBP1
ZFP36	SSB	TUT1	Ybx1	YBX1	TUT1	SFRS1
PCBP2	Sxl	HNRNPA1	CSDA	KHDRBS1	HNRNPA1	SFRS7
PCBP1	Sf1	SART3	ZFP36	SFRS1	gld-1	
ELAV1	KHSRP	ELAV1	SFRS1	HNRNPA1	Jsn1	
HNRNPK	MIR1236		IGF2BP1	ELAV1	Puf2	
RBM4	NCL		PCBP1	SLBP	Puf3	
U2AF2	asd-2		HNRNPK		Mpt5	
SFRS2	Msl5		HNRNPD		ACO1	
SFRS1			MEX3D		IREB2	
MBNL1			HNRNPA1			
Jsn1			ELAV1			
Puf2			PABPC1			
Puf3			TIA1			
Mpt5			TIAL1			
IREB2			SFRS9			
ACO1			HNRNPR			
			SYNCRIP			
			SLBP			
			SNRPB			
			NCL			
			GRSF1			
			Nab2			
			Puf4			
			Mpt5			

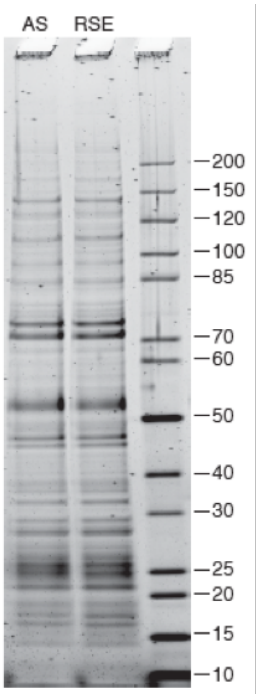
Table 1: Candidate proteins that bind highly mutated RSE regions

CCGGGU	GGUCUU	CUUACC	UUACUG	UUUGCG	GCGCGC	CGCUGU
NCL	PTBP1	PTBP1	PTBP1	Jsn1	PTBP1	PTBP1
gtf3a	ELAV1	PCBP1	CSDE1	Puf2	PTBP2	SFRS1
	TIA1	HNRNPK	HNRNPD	Puf3		SNRPA
	TIAL1	HNRNPA1	MEX3D	Mpt5		
	Slbp	SNRNP70	CPEB1			
		HNRNPD	KHSRP			
			FUBP1			
			ELAV3			
			PCBP1			
			PCBP2			
			NCL			
			PUM1			
			IGF2BP1			

Table 1 cont'd: Candidate proteins that bind highly mutated RSE regions. Shaded row header shows WT motif tested, with red nucleotides representing presence of mutations in that nucleotide position obtained from the mutagenesis screen. Columns are arranged in 5′ to 3′ order of mutation clusters. PTBP1 (green boxes) binds 12 of the 14 motifs tested. Other candidates include SFRS1 (blue boxes) and ELAV1 (orange boxes).

Efforts to identify proteins involved in RSV RSE stabilization were also concurrently made by our collaborators, Bobby Hogg and his postdoctoral fellow Zhiyun Ge, from the National Institute of Health (NIH). They showed that the RSV RSE is active in human cells, and can stabilize both viral transcripts and human reporter mRNAs with long 3′ UTRs when expressed in cis in heterologous constructs. [70, 71] Hogg and Ge used human reporter mRNAs in a PP7-based affinity purification and mass spectrometry screen to identify proteins that specifically bound the RSV RSE (see Figure 4). The human β-globin reporter mRNAs, containing the β-globin gene, SMG5 unstable long 3′ UTR, and the 400 nt RSV RSE versus an antisense RSE negative control, were used to pull down RNA-bound proteins. The comprehensive mass spectrometry screen yielded a number of proteins that were enriched in one sample over the other. PTBP1, which was enriched in the RSV RSE containing sample, was thus independently identified by the Hogg lab as a candidate protein that may play a role in the protection of long 3′ UTRs against NMD. In the same experiment, Upf1 protein is also underrepresented in the RSV RSE-containing sample as compared to the antisense control.





Protein ID	RSE	Antisense RSE
YB-1	133	196
Matrin 3	62	15
hnRNP L	36	17
ZAP	34	12
UPF1	27	45
PTBP1	17	3
SFRS14	13	0
RBM14	13	4
G3BP2	9	1
SART3	3	10

Figure 4: Identification of RSE-interacting proteins by tandem mass spectrometry. **Top:** Schematic of β-globin reporter mRNA constructs used for PP7-based affinity purification. The RSE sequence (top) and a control sequence, the antisense RSE sequence (bottom) were inserted into reporter mRNAs containing the β-globin gene and the SMG5 3′ UTR. **Left:** mRNP profiles of transcripts. Proteins were separated by SDS-PAGE and visualized by Krypton Infrared Protein Stain (Pierce). **Right:** Purified mRNPs were subjected to trypsin digestion and tandem mass spectrometry. Spectral counts from selected proteins enriched in either the RSE-containing sample or the antisense RSE sample are shown. Figure obtained from Ge and Hogg.

PTBP1 is a well-studied protein (reviewed by Romanelli et al) [72] best known for its role as a negative regulator of splicing. As its name suggests, it preferentially binds to CU-rich regions. It is thought that PTBP1 binds the polypyrimidine tract of the splice acceptor site, thus preventing spliceosome components from binding. In addition, PTBP1 contains four RNA recognition motifs (RRMs), which can bind different regions of the RNA and promote RNA looping, in line with its role as a splicing regulator. An analysis of the RSV RSE sequence shows that the element may contain up to eleven putative PTBP1 binding sites.

I verified this binding via an electrophoretic mobility shift assay (EMSA). Increasing amounts of PTBP1 protein were added to 40 ng of WT RSV RSE RNA and allowed to bind at 30ºC for 15 mins before the samples were run on 4% PAGE for 6 hrs at 4ºC. As shown in the top gel of Figure 5bi, PTBP1 binds WT RSE RNA and causes at least three discrete band shifts when added in increasing concentrations. The first shift starts to appear in lane 3, at 0.16:1 molecules of PTBP1 protein:WT RSE RNA. Lane 6, with 38.4 nM PTB, is where we see more RNA in the first complex than unbound RNA, and also the lane we first see a second 2nd shift. That corresponds to 1.21:1 molecules of PTBP1:WT RSE. The third shift is faint but observable in lane 7, and becomes clearer in lane 8. These discrete band shifts each indicate a specific complex formed by PTBP1 binding to the WT RSV RSE RNA. The binding did not result in large mobility shifts, which may mean that the RNA remained folded in the complex.

In lane 9, we can observe unbound WT RNA, the three different specific RSE-PTBP1 complexes, as well as a smear. One possible explanation of the smearing is

simply due to a supershift from additional proteins bound per RNA. The exact dynamics of PTBP1 binding to RNA are not well understood. Current understanding in the field suggests that stable PTBP1-RNA binding only requires 1 to 2 RRMs per PTBP1 molecule to be bound to a high-affinity binding site; a CU-dinucleotide will suffice for each of its other RRMs to bind. In addition, RRMs 3 and 4 can interact and bind to each other, both within the same molecule and across mutiple molecules to cause multimerization of the protein. Thus, at high protein concentrations when PTBP1 is present in excess of its binding substrates, smearing such as that observed in lane 9 may happen. In vast excess amounts of PTBP1, the protein and RNA aggregate, yielding a distinct band such as those observed in lanes 10–14 of the gels in Figure 5b.

However, a more likely explanation includes the unfolding of the RSE structure. As a negative control, I performed an EMSA using the 400 nt RSV antisense RSE RNA as a binding substrate for PTBP1 (Figure 5bi, bottom gel). Compared to the wildtype, I do not observe ribonucleoprotein (RNP) complex formations when the corresponding amounts of protein and AS-RSE RNA were used. Instead, smears were observed. This demonstrates the degree of non-specific RNA binding by PTBP1. Note that the smearing occurs in nearly the same concentration range as the specific complex formations in the WT RSE gel, but the change in mobility is large. This likely indicates a different type of complex formation, or a mixture of complexes with slightly different structures.

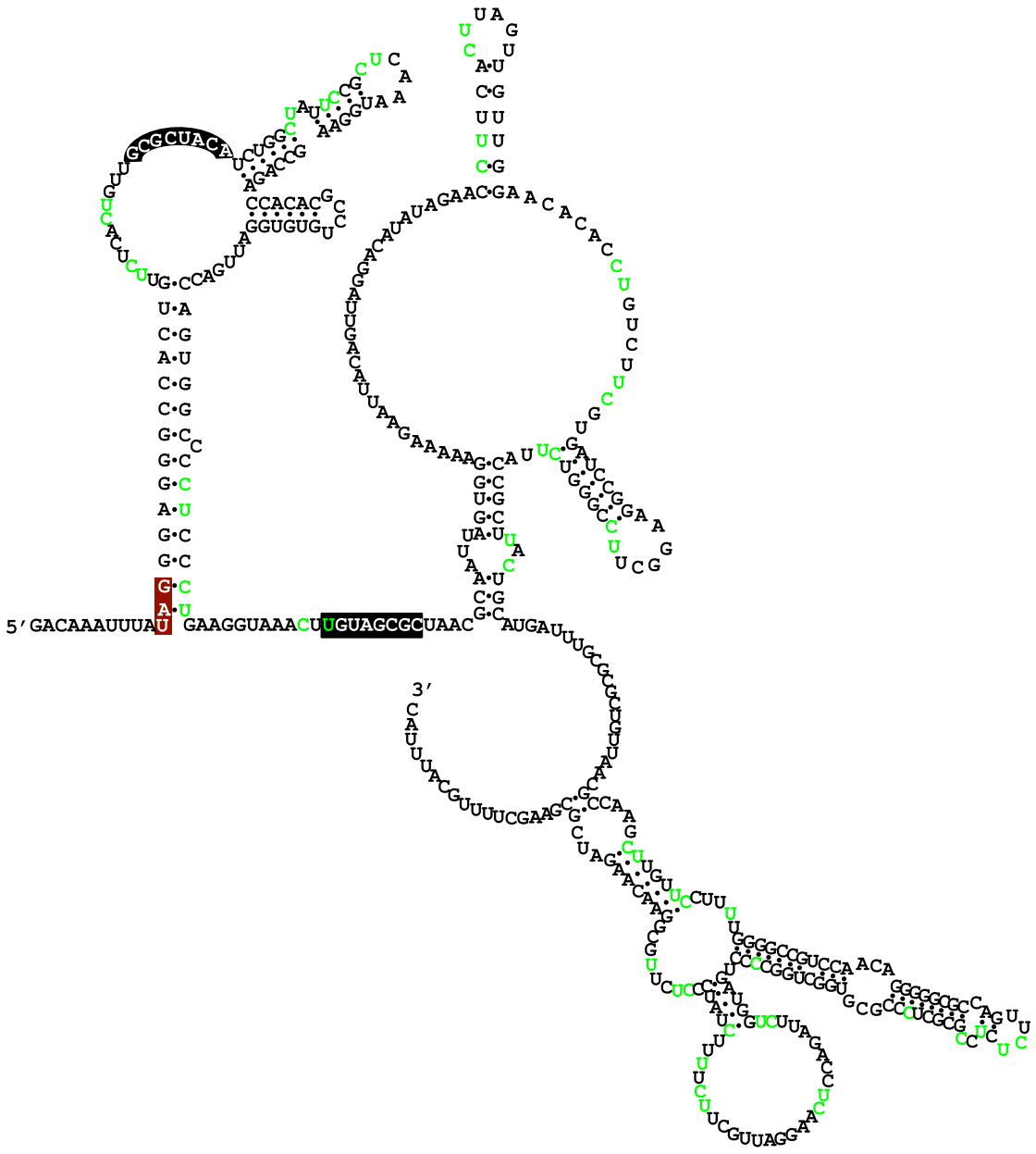


Figure 5(a)(i): WT RSV RSE. Nucleotides in green highlight CU-rich putative PTBP1 binding sites as predicted by sequence gazing.

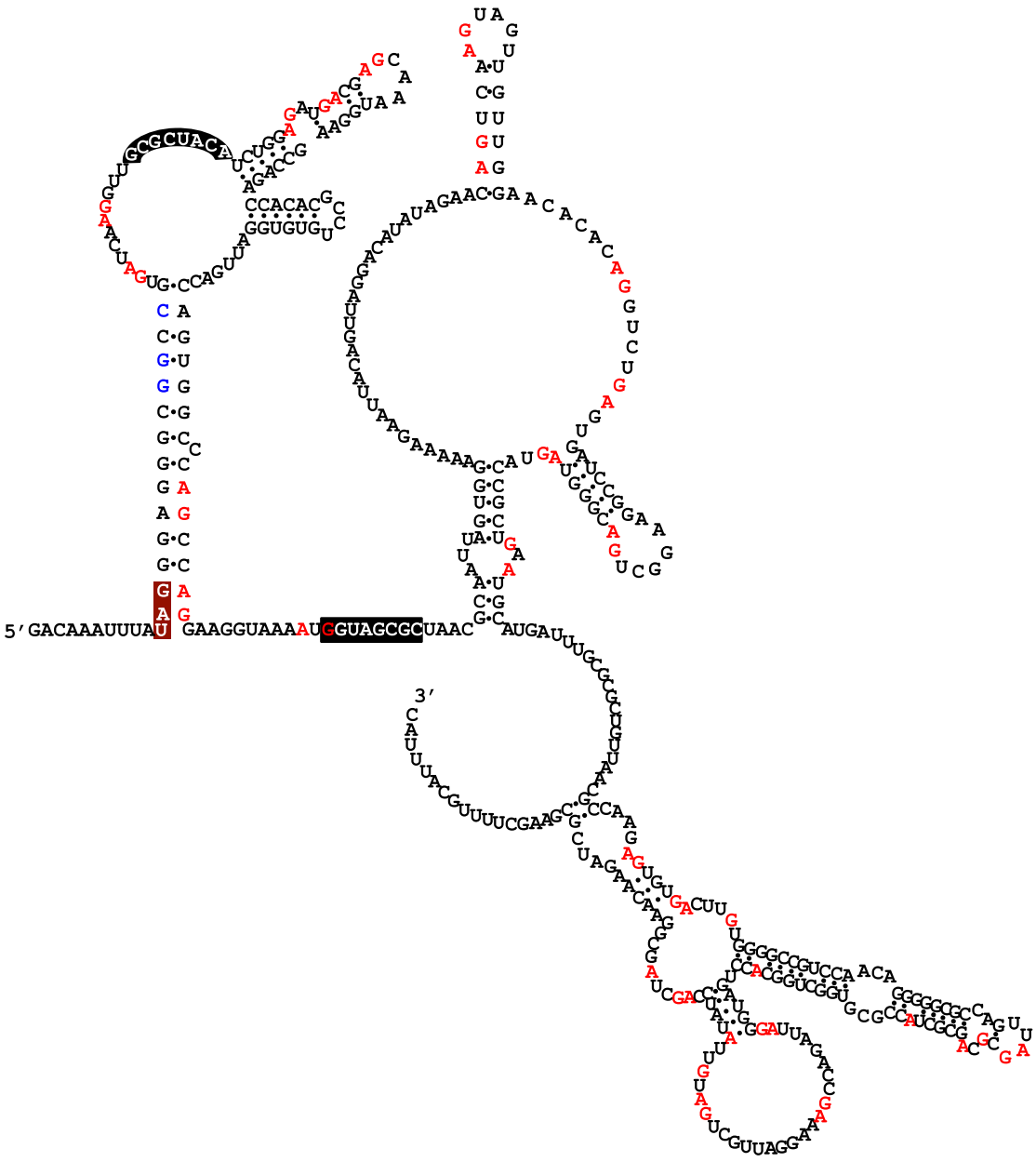


Figure 5(a)(ii): RSEΔPTB in RSV WT E/S vector. Nucleotides in red highlight nucleotides mutated to knock out putative PTBP1 binding sites. Cs were mutated to As, and Us were mutated to Gs. Nucleotides in blue represent the mutations made to generate the EagI site in the E/S vector.

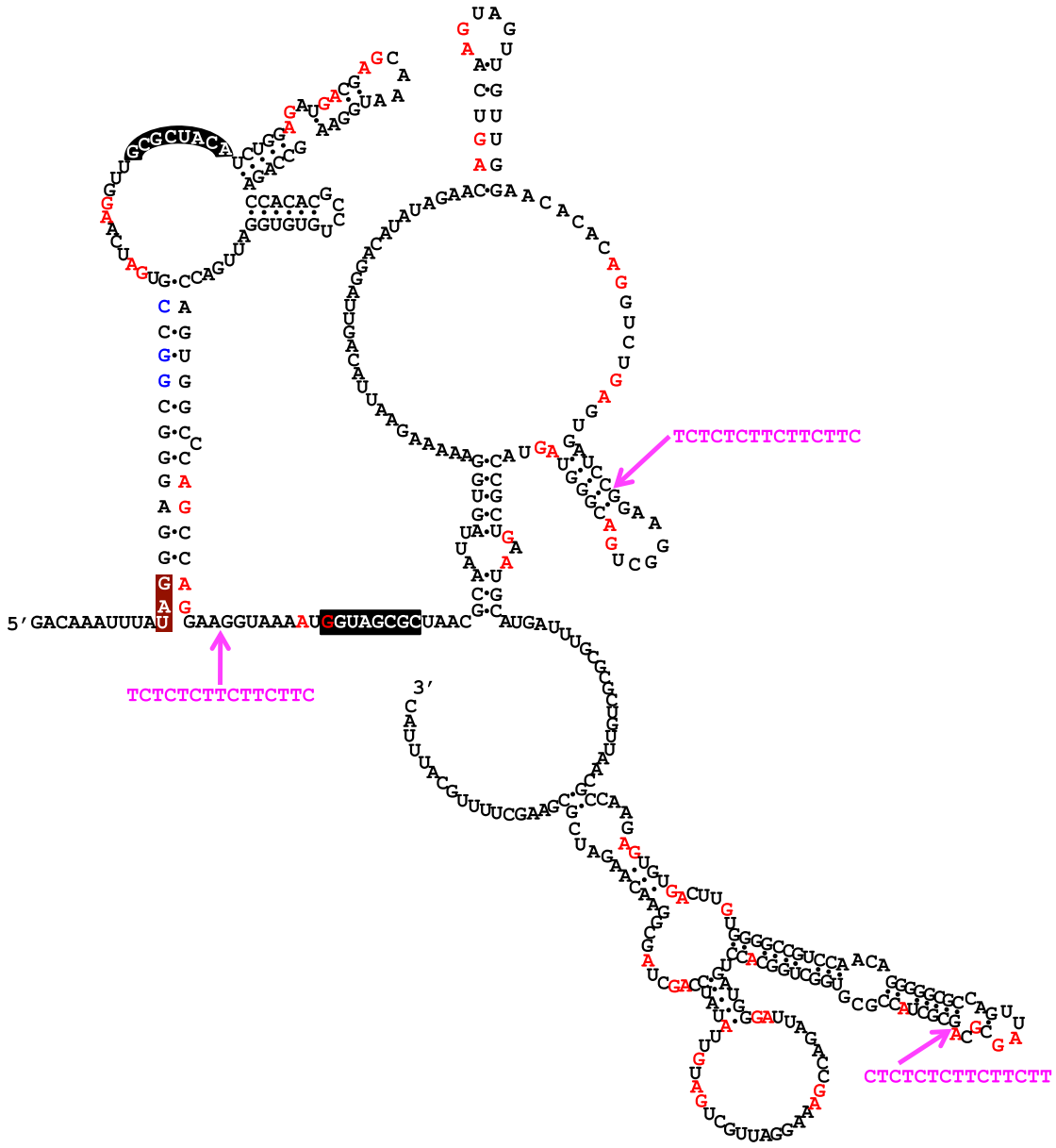


Figure 5(a)(iii): RSEΔPTB+3xPTB in RSV WT E/S vector. Nucleotides in red highlight nucleotides mutated to knock out putative PTBP1 binding sites. Cs were mutated to As, and Us were mutated to Gs. Nucleotides in blue represent the mutations made to generate the EagI site in the E/S vector. Arrows and nucleotides in magenta represent exogenous PTB binding sites added back to the RSEΔPTB construct.

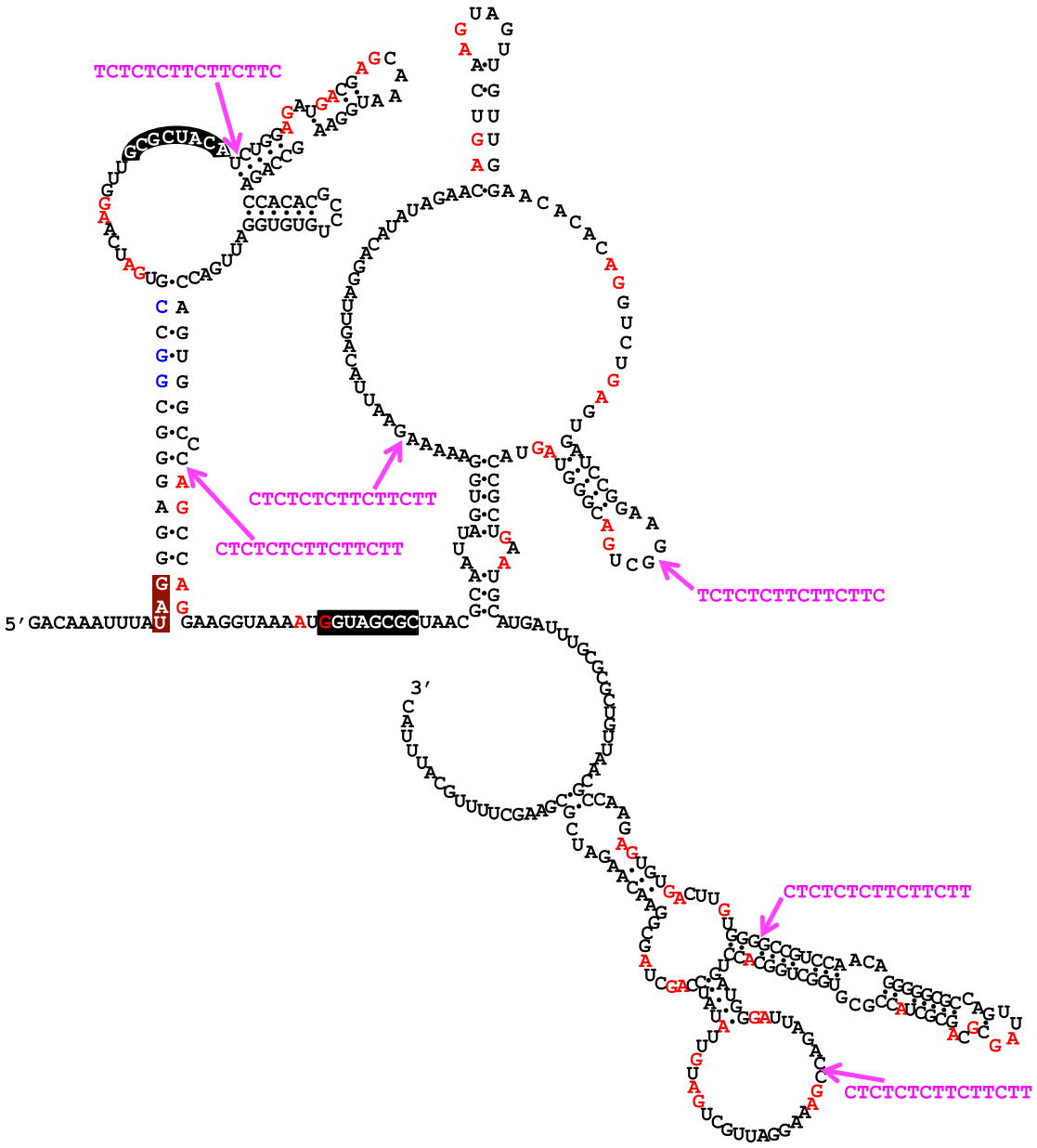


Figure 5(a)(iv): RSEΔPTB+6xPTB in RSV WT E/S vector. Nucleotides in red highlight nucleotides mutated to knock out putative PTBP1 binding sites. Cs were mutated to As, and Us were mutated to Gs. Nucleotides in blue represent the mutations made to generate the EagI site in the E/S vector. Arrows and nucleotides in magenta represent exogenous PTB binding sites added back to the RSEΔPTB construct.

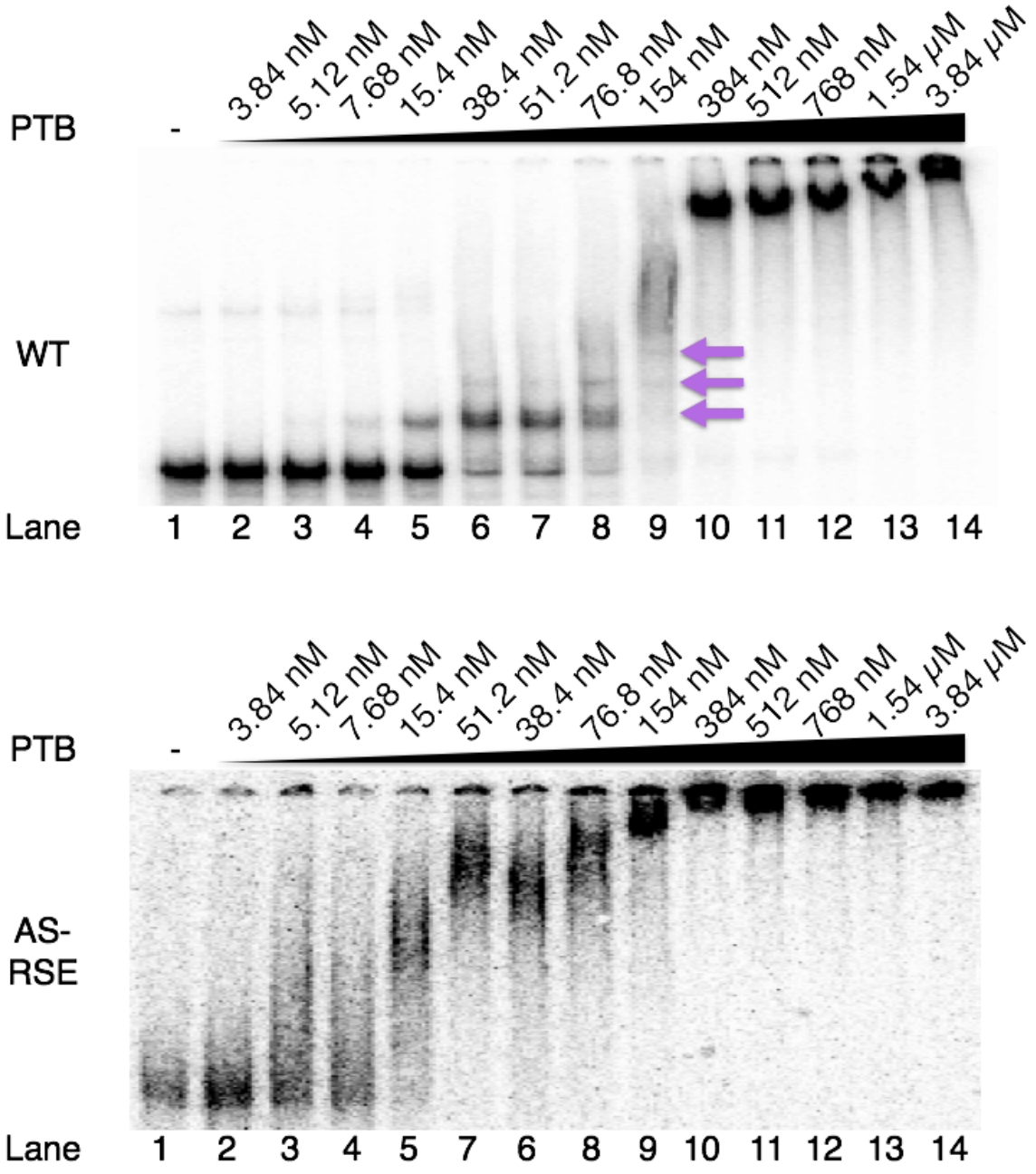


Figure 5b(i): EMSAs of RSV RSE variant RNAs and PTBP1 protein. Purple arrows mark RSE-PTB complex formations. The 400 nt WT RSV RSE RNA forms at least 3 different complexes with PTBP1. Complex formation is lost with the 400 nt AS-RSE RNA (samples 6 and 7 were inadvertently switched and loaded in the wrong lanes).

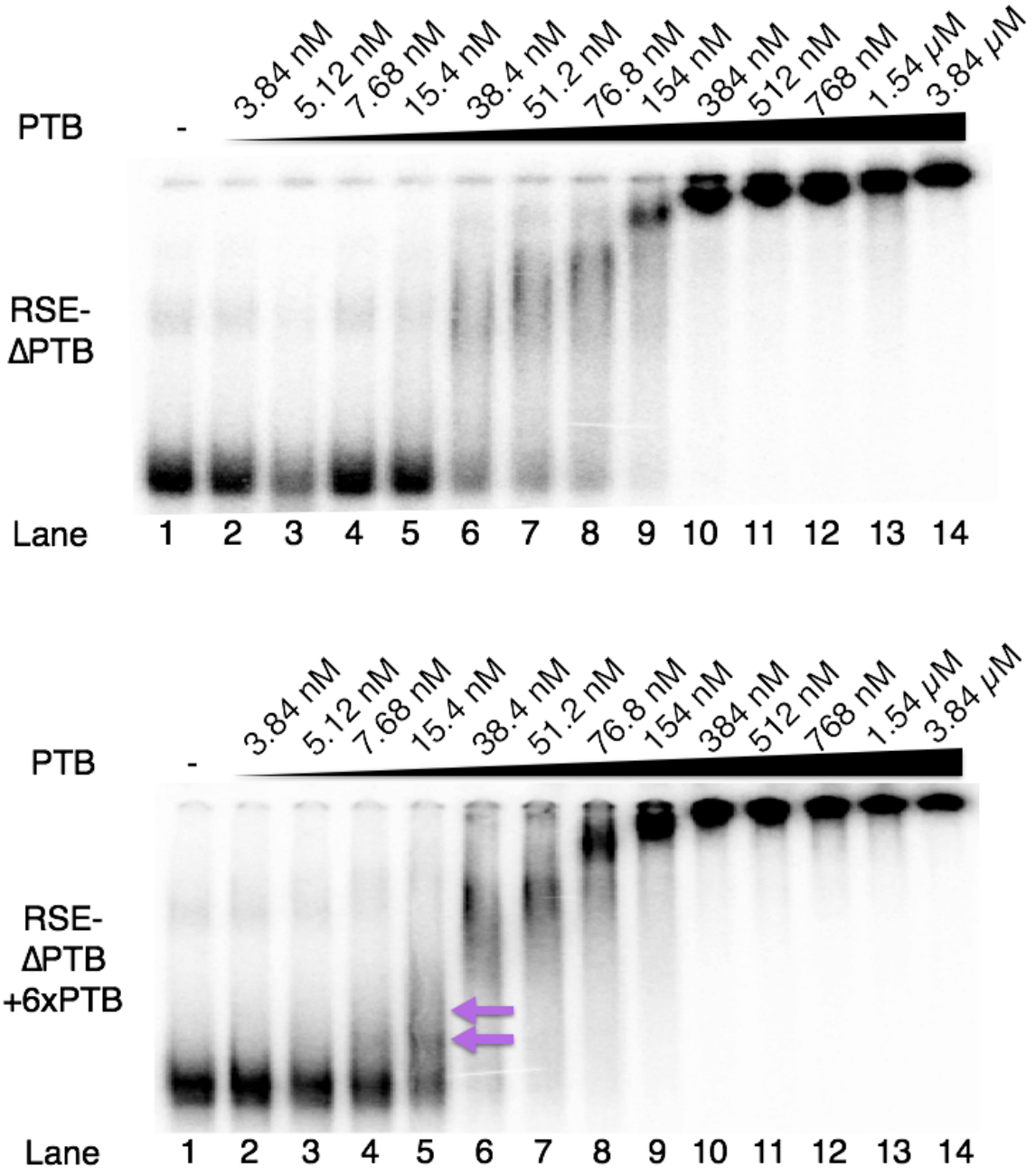


Figure 5b(ii): EMSAs of RSV RSE variant RNAs to PTBP1 protein. Purple arrows mark RSE-PTB complex formations. Complex formation is lost with the 400 nt RSEΔPTB RNA. RSE-PTB binding is partially restored with the 490 nt RSEΔPTB+6xPTB RNA.

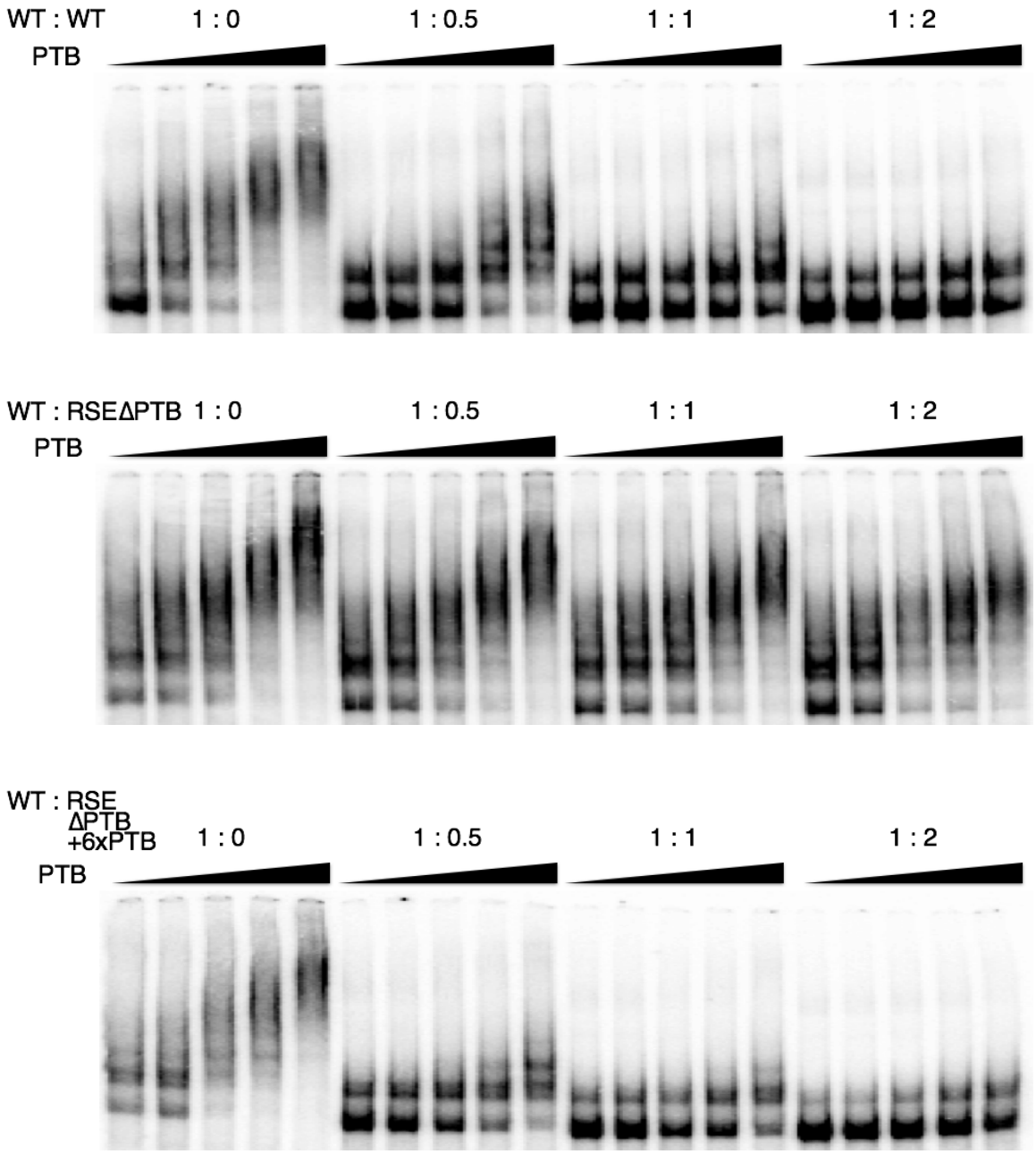


Figure 5c: Hot WT RSV RSE RNAs versus cold RSV RSE variant RNAs gel shift competitions. Cold WT RSE RNA efficiently compete for PTB binding against hot WT RSE RNA, resulting in decreased hot WT RSE-PTB complex formation with increasing cold WT RSE RNA added. RSEΔPTB RNA is unable to compete with WT RSE RNA, demonstrating poor RNA-protein binding. RSEΔPTB+6xPTB is an efficient competitor.

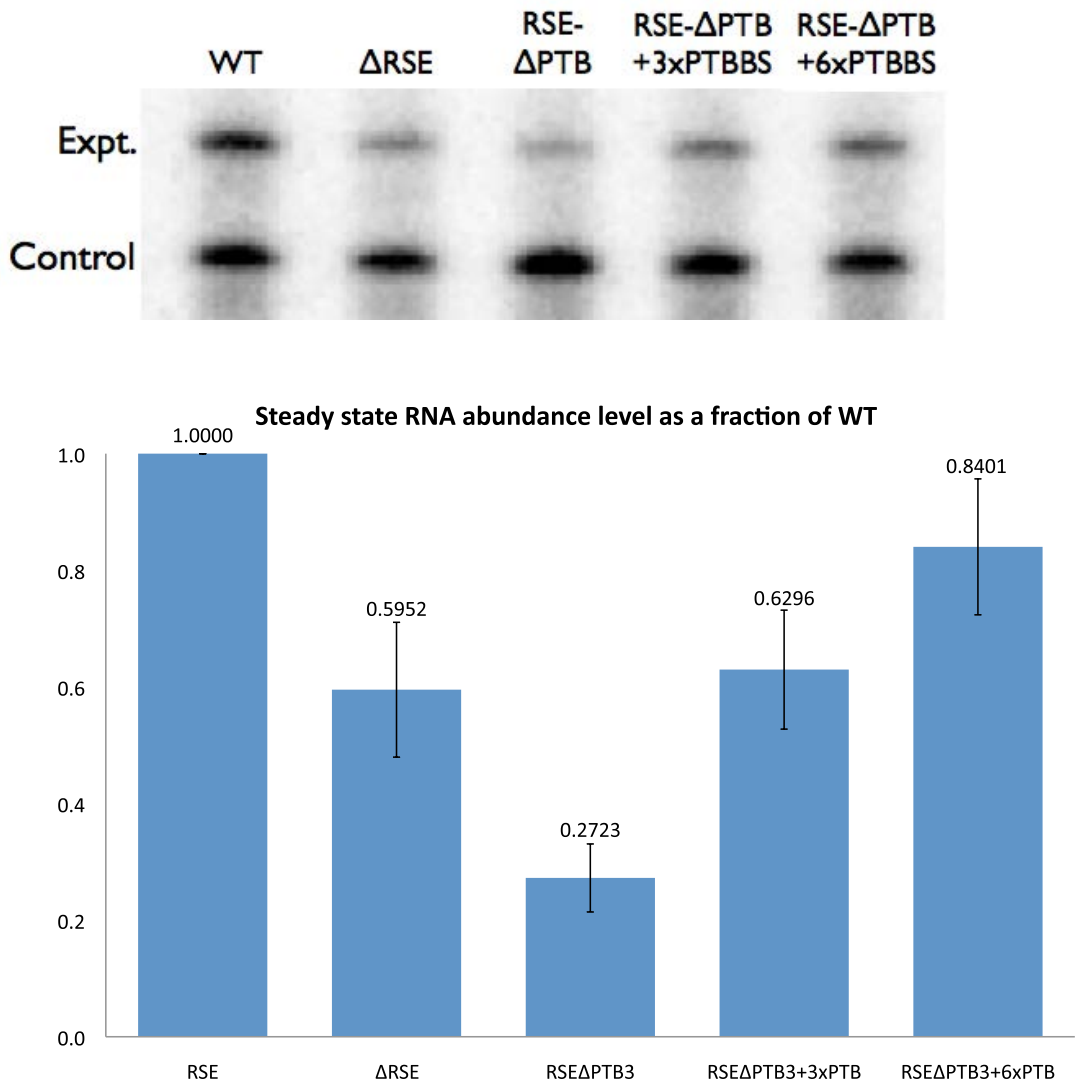


Figure 5d: PTBP1 binding sites directly affect the stability of unspliced RSV RNA. **Top:** Representative gel of an RPA on RSV unspliced RNA containing RSE variants. **Bottom:** Quantification of RPAs. Levels of the experimental unspliced RSV RNAs (experimental band) were normalized to levels of the transfection control (control band). RNA levels are reported as a fraction of RSV RNA containing WT RSE. Error bars indicate ± SD; n = 5, $p < 0.05$

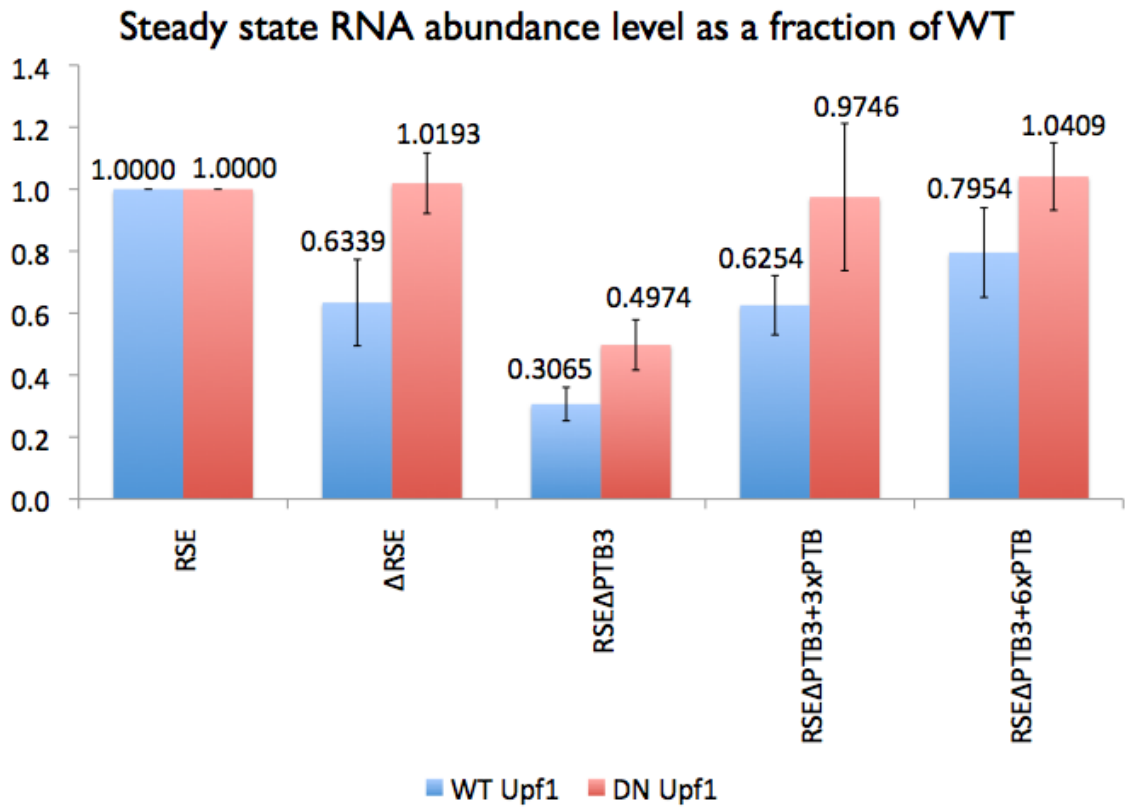


Figure 5e: RSV unspliced RNA abundance levels can be rescued by cotransfecting dominant negative Upf1. Quantification of RPAs. RNA levels are reported as a fraction of their respective WT RSE RNA; samples co-transfected with wildtype Upf1 plasmid (blue) were normalized to the blue WT RSE sample, and samples co-transfected with dominant negative Upf1 plasmid (red) were normalized to the red WT RSE sample. Error bars indicate ± SD; n = 5

PTBP1 is important for RSE-mediated RNA stability

To test the role of PTBP1 in stabilizing long 3′ UTRs, Hogg and Ge generated a ΔPTB mutant, in which they mutated all putative PTBP1 binding sites in RSV RSE by mutating CU dinucleotides to AGs (see Figure 5aii versus 5ai). Two additional RSV RSE variants were also generated, in which three and six exogenous PTBP1 binding sites that had been previously validated were added back to the ΔPTB mutant (Figure 5aiii and 5aiv respectively). This allows us to rescue any potential phenotype the ΔPTB mutant may exhibit.

EMSAs were performed to evaluate PTBP1 protein binding to these RSV RSE variant RNAs (see Figure 5bii). With RSEΔPTB, I observed smeared bands in the same concentration range as the non-specific binding to AS-RSE in Figure 5bi. This smearing was partially rescued by the addition of 6 PTBP1 binding sites in the RSEΔPTB+6xPTB variant. This suggests that WT RSE RNA forms specific, structured complexes with PTBP1 whereas the AS-RSE, RSEΔPTB, and to a lesser extent the RSEΔPTB+6xPTB, binds PTBP1 non-specifically. The difference in binding affinity does not appear to be large; the main differences between the WT RSE and the RSE variants appear to lie in the structural homogeneity and compactness of the specific versus non-specific complexes.

In addition, I performed gel shift competitions, using cold RSE variant RNAs as competitors for radiolabeled WT RSE RNA (see Figure 5c). 40 ng of hot WT RSE RNA were used to bind PTBP1 at concentrations ranging from 7.68 nM to 76.8 nM, corresponding to lanes 4–8 of the EMSAs in Fig 5b. Cold RSE variant RNAs were then added in different ratios. Cold RSEΔPTB RNA was not an efficient competitor for hot

WT RSE RNA, showing that it does not bind PTBP1 as well as WT RSE RNA. On the other hand, cold RSEΔPTB+6xPTB RNA was able to compete with WT RSE RNA and prevent RNP complex formation, showing that the RNA is able to bind PTBP1.

I cloned the RSV RSE variants back into the full-length virus and performed RPAs to evaluate the effects of PTBP1 binding on full-length viral RNA abundance (see Figure 5c). ΔPTB shows a very striking decrease in abundance levels of RSV RNA; this mutant is present in even lower levels than deleting the entire 400 nt RSE. The addition of three and six PTBP1 binding sites back to the ΔPTB mutant does rescue RNA levels in a dose-dependent fashion. Furthermore, co-transfection of dominant-negative (DN) Upf1 restores RNA levels in the addback mutants back to WT levels (see Figure 5d). This tells us that these RSE sequences, which are putative PTBP1 binding sites, are important for protecting full-length RSV RNA against NMD.

PTBP1 binds RSV RSE to prevent Upf1 binding

Hogg and Ge also quantitated PTBP1 versus Upf1 protein binding to these RSE variants. Using PTBP1 and Upf1 in a co-immunoprecipitation experiment, they pulled down human reporter mRNAs from cell lysates and performed northern blots to detect the levels of the respective RSE-variant-containing RNAs recovered. The results are shown in Figure 6, where we can observe an anti-correlative trend. Comparing the ΔPTB to the 6xPTB, it is clear that when there is more PTBP1 binding, we correspondingly see less Upf1 binding.

From these data, we developed a model of the mechanism of RSV RSE protection against NMD (refer to Figure 7). The RSV RSE binds PTBP1, which competitively inhibits Upf1 from binding downstream of the gag stop codon. As such, even though the RNA still has a long 3′ UTR when the ribosome terminates, the ribosome is unable to sense the actual 3′ UTR length since there is no downstream Upf1 molecules to interact with the ribosome to trigger NMD.

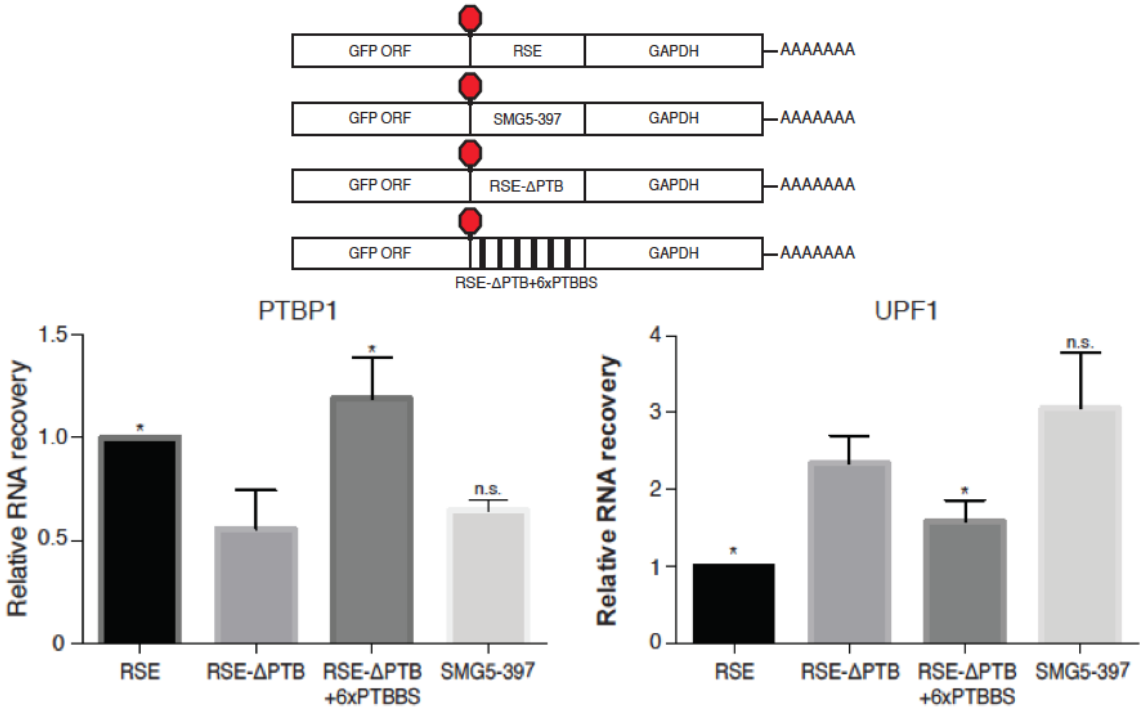


Figure 6: Accumulation of PTBP1 on the 3′UTR prevents UPF1 binding. Top: Schematic of reporter mRNA constructs used in immunoprecipitation assays. **Left**: Quantification of relative RNA recovery upon PTBP1 immunoprecipitation, normalized to the recovery of co-transfected GFP-bGH mRNAs. PTBP1 is reduced on RSEΔPTB transcripts as compared to WT RSE and RSEΔPTB+6xPTBBS. **Right**: Quantification of relative RNA recovery upon UPF1 immunopreciptation, normalized to the recovery of co-transfected GFP-bGH mRNAs. Upf1 is increased on RSEΔPTB transcripts as compared to WT RSE and RSEΔPTB+6xPTBBS. Error bars indicate ± SD; $n \geq 3$ (*: $P < 0.05$ in two-tailed Student's t-tests when compared to RSE-ΔPTB recovery). Data obtained from Ge and Hogg.

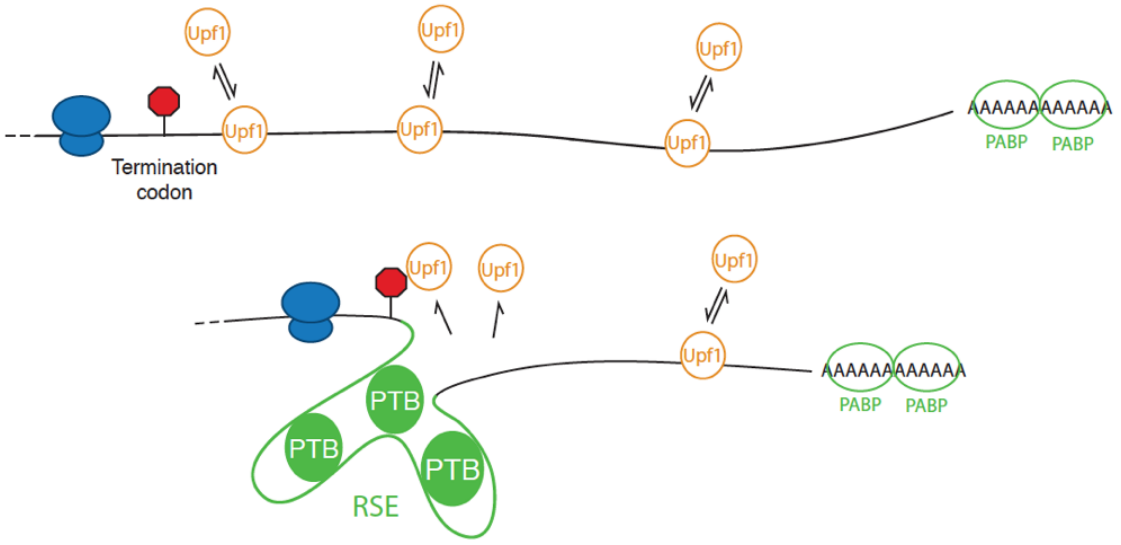


Figure 7: **RSV RSE binds PTBP1 to protect RSV RNA from NMD.** In the absence of the RSE or other sequences capable of recruiting PTBP1, UPF1 binds 3′UTRs in a length-dependent manner, potentiating NMD. PTBP1 can bind the RSE at multiple sites throughout the 400 nt sequence. This establishes a higher-order mRNP structure in the vicinity of the stop codon that is refractory to UPF1 binding. The NMD pathway thus judges the 3′UTR to be short and fails to degrade the mRNA.

PTBP1 binding sites are enriched downstream of the gag stop codon in RSV genome

Why does the ΔPTB mutant show a more severe phenotype than deleting the entire RSE? To answer this question, I looked at PTB binding site distribution across the entire RSV genome. Hogg and Ge had previously analyzed and identified a list of pentamer and hexamer sequences that were enriched in publicly available PTB CLIP-seq data sets, which signify high-affinity PTBP1 binding (see Table 2). Using a python script written by Nathan Lee, an undergraduate research assistant in the Beemon lab, I scanned the RSV genome for these high-affinity binding sites and plotted their frequency in 50 nt bins (see Figure 8a).

Pentamers	Hexamers
UCUCU	UUCUCU
UCUGU	UCUUCU
CUUCU	UCUCUU
UCUUC	UCUCUG
UUCUC	CUUUCU
CUUUC	CUUCUC
UUUCU	UCUCUC
UUCUG	UUCUGU
UGUCU	CUCUCU
UUCCU	CUGUCU
CUCUC	UUUCUC
UCUUU	UUUCCU
UUCUU	UUUCUG
CUCUU	UCUUUC
CUGUC	GUCUCU
UCCUU	UGUCUU
UAUCU	CUUUUC
CUCUG	UUCUUC
UCUAU	CUUUCC
GUCUU	CUCUUU

Table 2: Over-represented motifs in PTB CLIP experiments

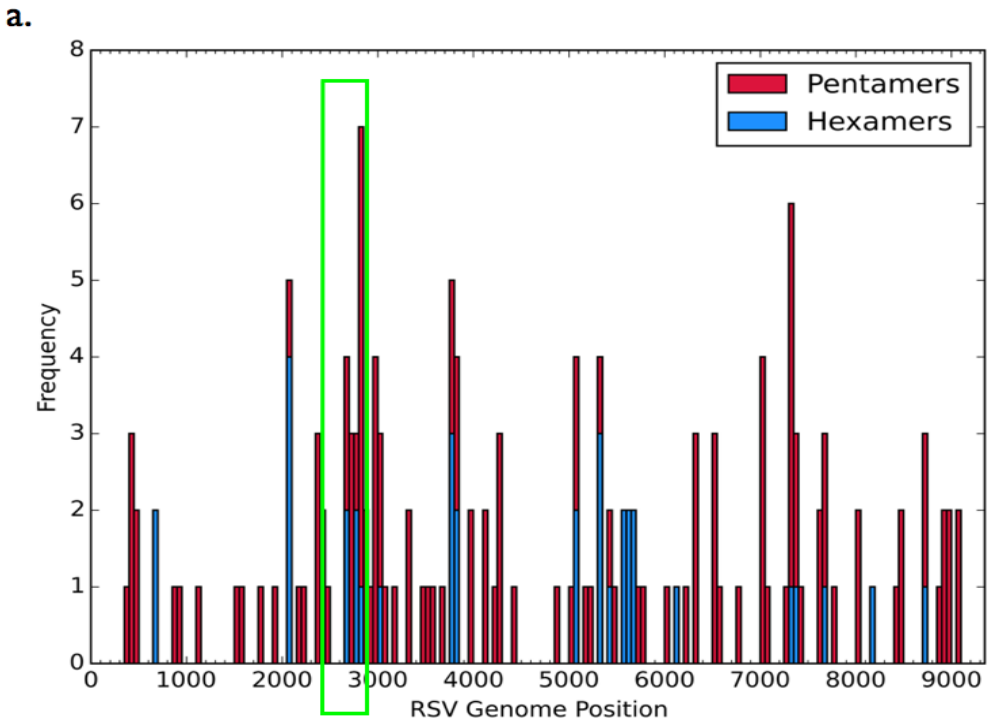


Figure 8(a): PTB binding site distribution across RSV genome. PTB binding sites are most enriched downstream of gag stop codon. RSV RSE marked by green box.

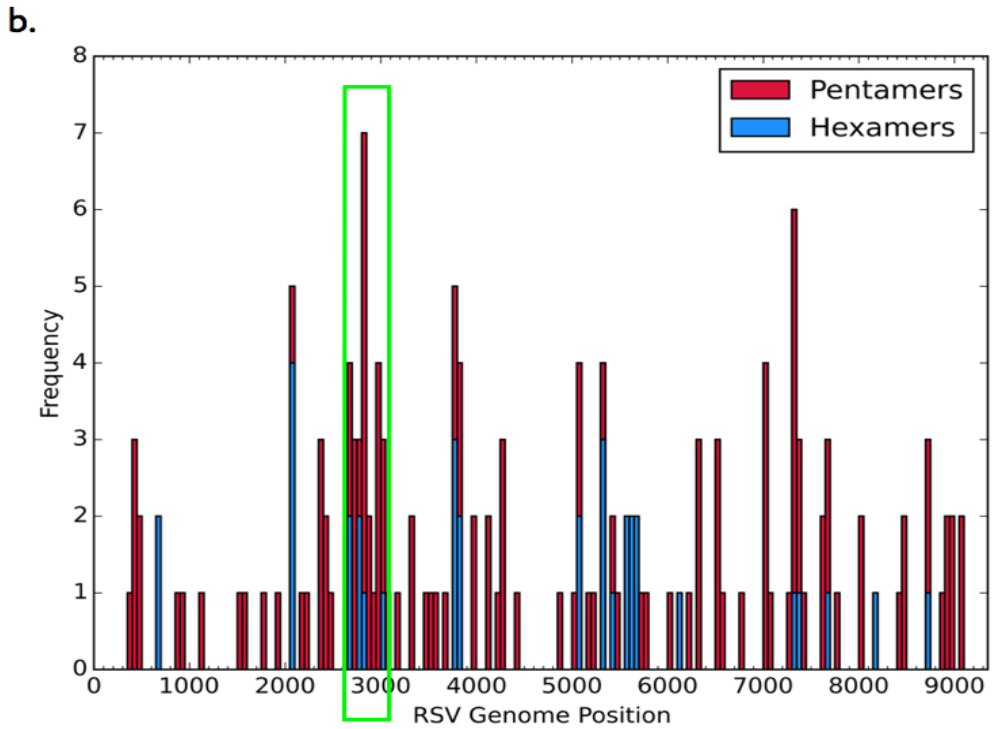


Figure 8(b): Redesignation of the RSV RSE. Green box marks highest concentration of PTB binding sites across RSV genome, and suggests that this region (nts 2650–3050) represents the true RSV RSE element that would be most effective in protecting the full-length transcript from NMD.

PTBP1 binding sites are most enriched in the region downstream of the gag stop codon, including the RSE (green box in Figure 8a). The 5′ end of the box shows a region devoid of PTBP1 binding sites; this region corresponds to the ribosomal frameshifting pseudoknot, and corroborates our previous finding that deletion of the pseudoknot does not affect RNA stability. In addition, there are a number of PTBP1 binding sites immediately downstream of the box, which explains why ΔPTB has a more severe effect than ΔRSE: When the entire RSE is deleted, I am essentially shifting those downstream PTBP1 binding sites closer to the RSV gag stop codon, hence yielding a partial phenotype. In contrast, only the PTBP1 binding sites are mutated in the ΔPTB mutant. This creates a zone of PTBP1 exclusion, in turn allowing Upf1 to bind and trigger NMD of the RNA. In hindsight, the 400 nt region of the RSV 3′ UTR designated as the RSV RSE is an arbitrary attribution. In light of this PTBP1 binding site distribution data, we should adjust the designation of the RSV RSE to reflect the region encompassing the maximum number of PTBP1 binding sites, as shown by the box in Figure 8b, which spans nts 2650-3050.

Structural analysis of the RSV RSE RNA

A closer analysis of Figure 6 indicates that we still do not fully understand how the RSV RSE functions. Even though the 6xPTB RSE variant binds PTBP1 better than WT RSE, the WT RSE is still superior in inhibiting Upf1 binding. There are a few possible explanations for this, including RNA structure or cooperative binding with other proteins.

RNA structure may directly affect RNA stability by way of affecting PTBP1 positioning with regards to the stop codon. For instance, the 5′ end of the minRSE and C-frag is a CU-rich region that is 92 nts away from the gag stop codon. However, due to the secondary structure of the RNA and the folding of the ribosomal frameshifting pseudoknot, that region actually sits directly next to the stop codon at the base of the stemloop, and may play a bigger role in RNA stability than otherwise suggested. Likewise, the rest of the WT RSV RSE may fold in a way that could either maximize PTBP1 binding, or adopt a conformation that would best occlude Upf1 from the region and from the terminating ribosome.

In support of this hypothesis, we performed Mfold on the RSV RSE variant RNAs (see Figure 9). The Mfold software predicts optimal and suboptimal 2º structures for RNA and DNA molecules based on minimum free energy. [73] The predicted 2º structures of all 3 RSE variants are severely disrupted and differ significantly from the WT RSV RSE structure. Whereas the WT RSV RSE has 3 distinct domains, consisting of the 5′ frameshift pseudoknot, a 3′ GC-rich stemloop, and a flexible central region, the RSEΔPTB and RSEΔPTB+3xPTB both form a large, single stemloop, while the RSEΔPTB+6xPTB adopts a Y-shaped conformation.

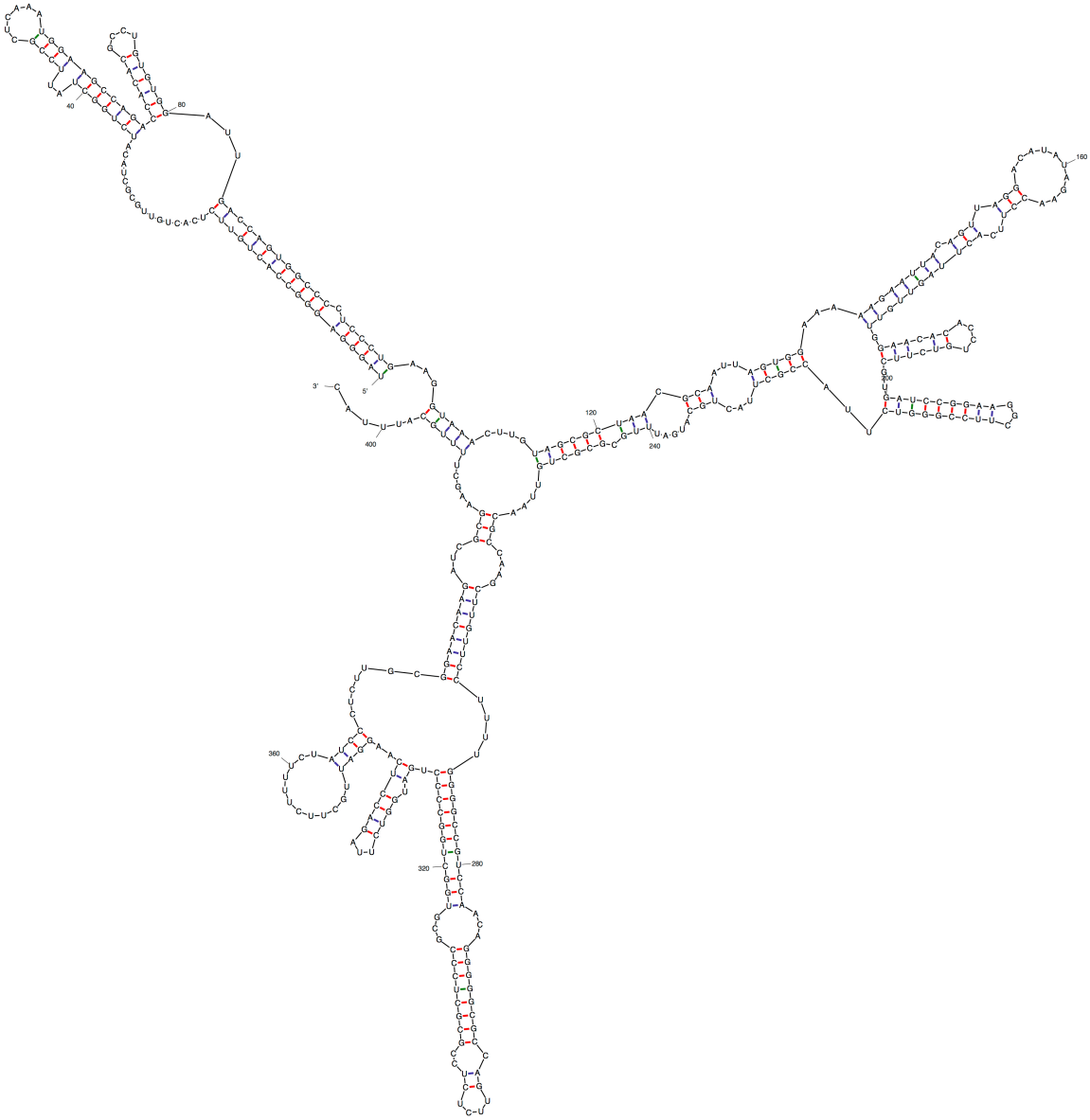


Figure 9: Mfold of WT RSV RSE variant RNA. Above: WT RSV RSE RNA. **On next page:** RSEΔPTB and RSEΔPTB+3xPTB. **On page 48:** RSRΔPTB+6xPTB. Nucleotides in blue represent the mutations made to generate the EagI site in the E/S vector. Nucleotides in red correspond to C to A and U to G mutations made to knock out putative PTBP1 binding sites. Nucleotides in magenta represent exogenous PTBP1 binding sites added back to the RSEΔPTB construct.

Figure 9: Mfold of RSV RSE variants. Left: RSEΔPTB. **Right:** RSEΔPTB+3xPTB

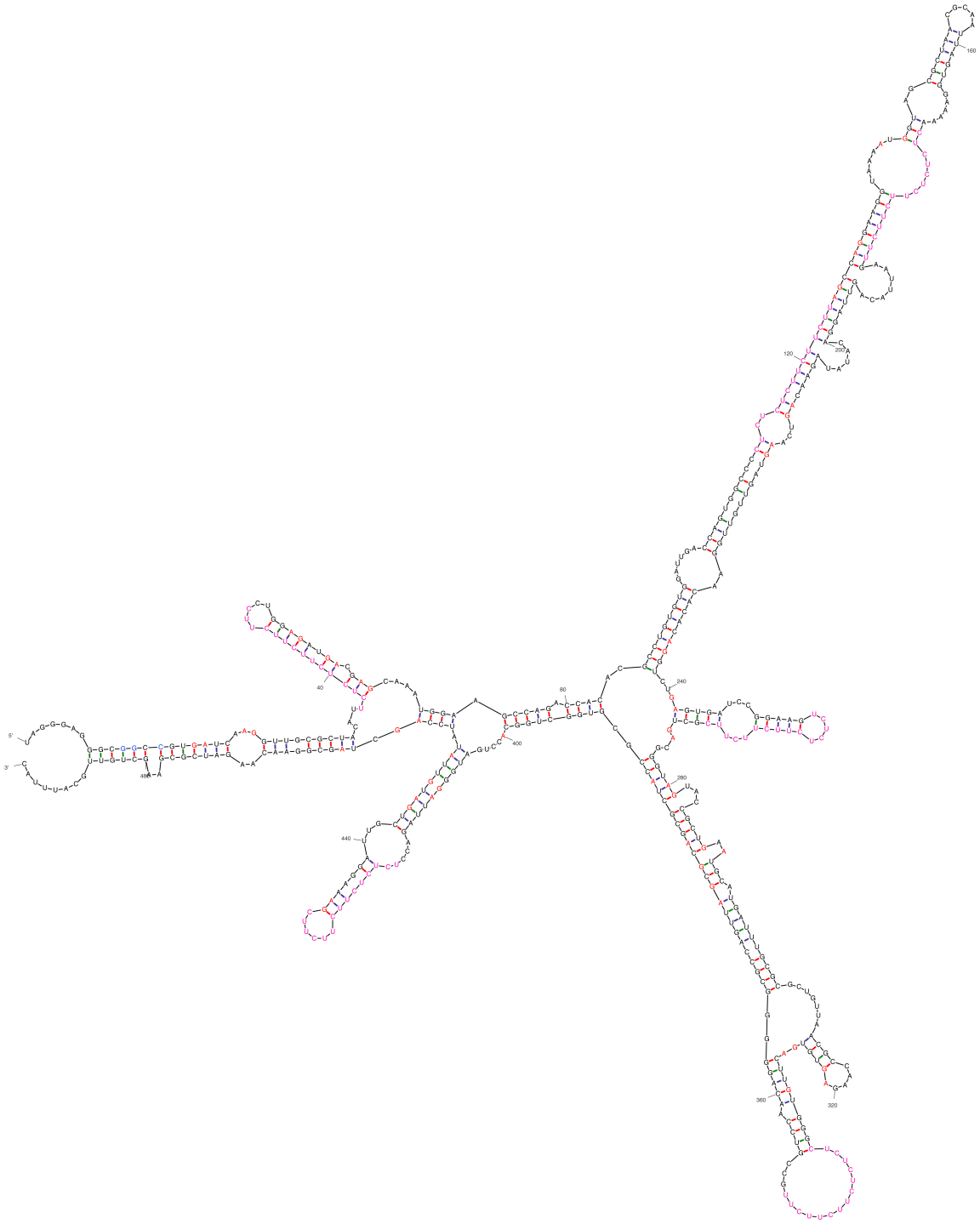


Figure 9: Mfold of RSV RSE variants. Above: RSEΔPTB+6xPTB.

Future directions

To pursue this further, we initiated two collaborations with structural RNA labs. The first collaboration is with Yuan Yuan Liu from the Michael Summers Lab at University of Maryland – Baltimore County. The Summers Lab is renowned for pushing the boundaries of RNA size limitations for 2-dimensional nuclear magnetic resonance (2D-NMR). For this collaboration, I visited the Summers Lab and carried out large-scale synthesis and purification of RNAs (see Figure 10a). WT RSV RSE RNA appears to be well behaved; both the minRSE RNA and the C-frag RNA run as a single species on a native gel in H_2O and in isoelectric PI buffer (refer to Figure 10b). Additionally, a diagnostic 1D-NMR performed on the minRSE RNA shows that the RNA yields the same spectra, with or without salt, with sharp and distinct peaks that can be resolved and assigned, making it an ideal candidate for 2D-NMR (see Figure 10c). As such, the Summers Lab is proceeding with the 2D-NMR experiment on the minRSE RNA. 2D-NMR on the C-frag RSE is much more technically challenging due to size limitations, but we may still pursue it by performing the NMR on partial fragments.

The second collaboration we initiated is with Rossitza Irobalieva from the Wah Chiu Lab at Baylor College of Medicine to perform cryo-electron microscopy (cryo-EM) on the RSE RNA. Figure 10d shows preliminary results obtained on the cryo-EM of the minRSE. minRSE RNA appears to adopt four different conformations, suggesting structural flexibility within the RNA. As cryo-EM is easier with larger RNAs, the Chiu Lab is proceeding with cryo-EM of the C-frag RSE, the full 400 nts of the WT RSV RSE, as well as protein-RNA complexes of PTBP1 in combination with these RNAs.

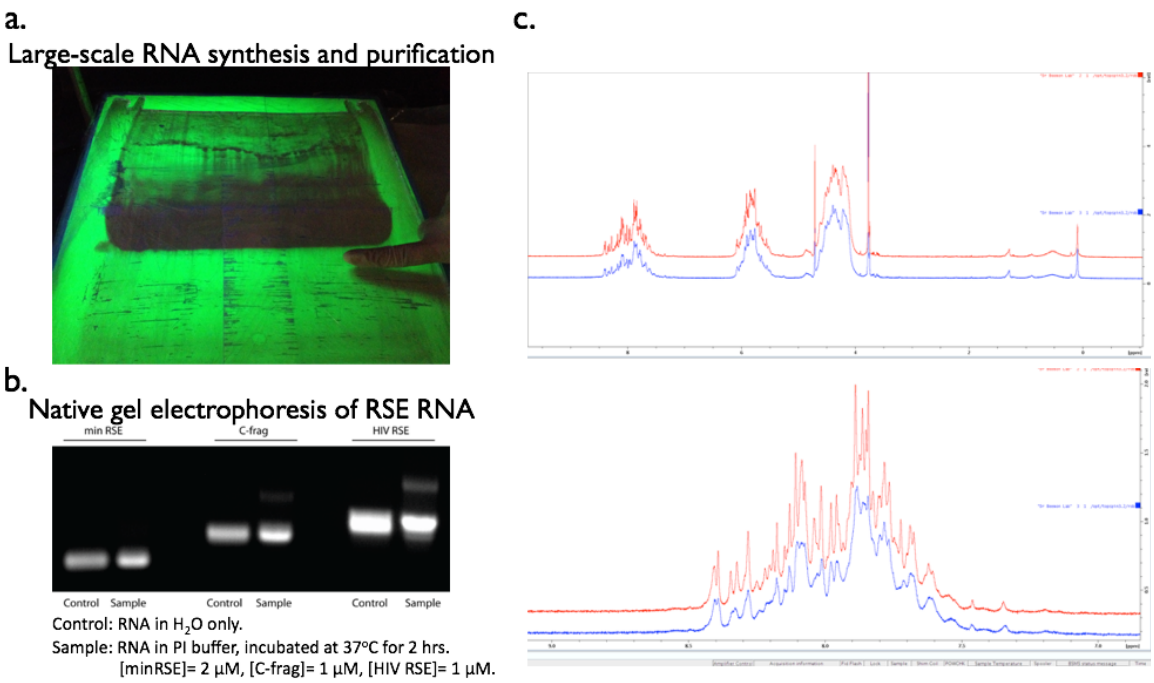


Figure 10: Preliminary RSV RSE RNA structural data from collaborations. (a) Large-scale RNA synthesis using in vitro transcription with T7 RNA polymerase. **(b)** Native gel electrophoresis shows that minRSE, C-frag RSE and HIV putative RSE run as a single species both in H_2O and in isoelectric point buffer. **(c)** 1D-NMR on minRSE RNA yields the same spectra with (red trace) and without (blue trace) $MgCl_2$.

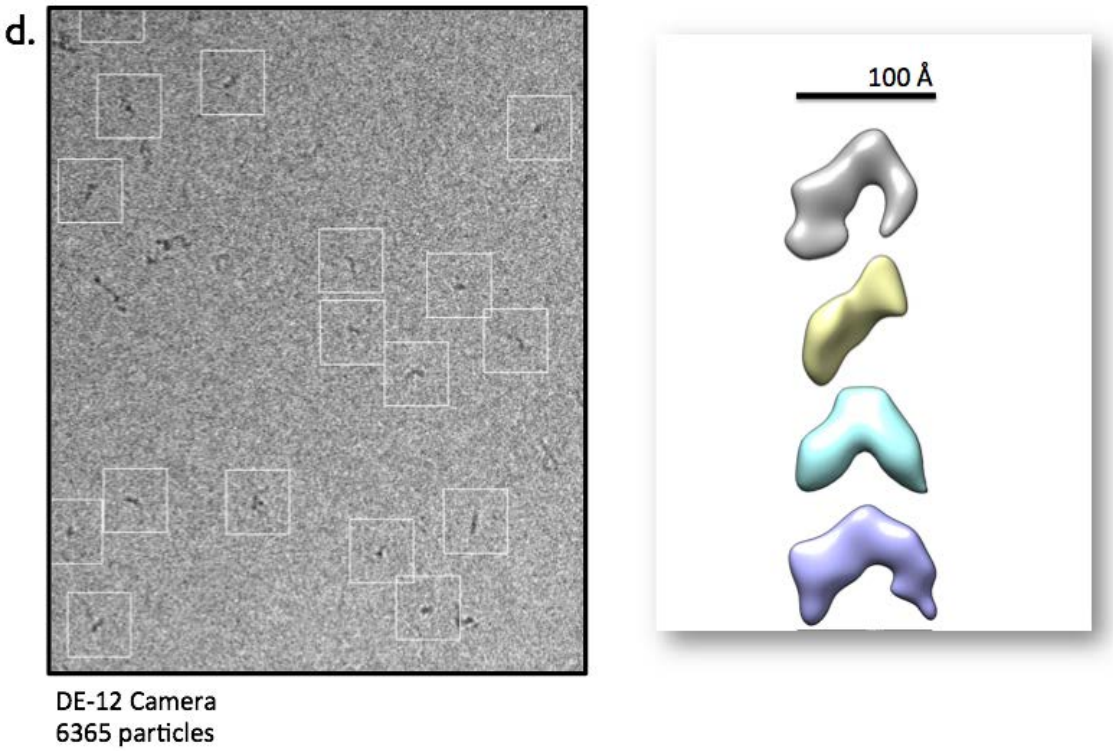


Figure 10: Preliminary RSV RSE RNA structural data from collaborations. (d) minRSE RNA exhibits four structural classes of conformation under cryo-EM, suggesting that it is structurally flexible.

CHAPTER 3:

INVESTIGATING PUTATIVE RNA STABILITY ELEMENTS IN HUMAN IMMUNODEFICIENCY VIRUS (HIV) AND MOLONEY-MURINE LEUKEMIA VIRUS (MMLV)

Background

Are similar stability elements present in other retroviruses, such as HIV and MMLV? As can be seen in Table 3, the features that make RSV prone to degradation by NMD are not unique traits, but rather traits that are shared by all retroviruses. These viruses have very similar genome sizes and structure, generating polycistronic mRNAs that code for multiple genes, including gag, pol and env. Likewise, other viruses that contain multiple open reading frames may also induce NMD.

For HIV and MMLV, since the translating ribosome terminates at the gag stop codon the majority of the time, the long 3′ UTR of the full-length unspliced RNA from these viruses should trigger NMD, causing the rapid decay of their RNAs. As viruses are unable to modulate their 3′ UTR lengths, it is plausible that these other retroviruses may also have evolved RSE-like stability elements that would help protect their long 3′ UTRs from NMD.

Feature	RSV	MMLV	HIV-1
Genome size (kb)	9.3	8.3	9.2
Genes	gag, pol, env, src	gag, pol, env	gag, pol, env, vif, vpr, vpu, nef, tat, rev
Long 3´UTR	yes	yes	yes

Table 3: MMLV and HIV-1 share features of RSV that make them susceptible to NMD. RSV and MMLV are simple retroviruses whereas HIV is a complex retrovirus. However, they have similar genome sizes and structures, and their unspliced RNAs have long 3´ UTRs that may trigger NMD.

400 nts of HIV 3′ UTR immediately downstream of the gag stop codon can stabilize full-length RSV RNA in a chimeric construct.

Since the tools and protocol for evaluating RNA stability in RSV have already been developed and optimized, I first wanted to test the potential of the 3′ UTR immediately downstream of the HIV gag stop codon in stabilizing RSV RNA. To do this, I generated an HIV-RSV chimeric virus by cloning the 400 nt region of the HIV 3′ UTR (nts 2303–2692) into RSVΔRSE (see Figure 11a). Additionally, I generated 15 other variants of the HIV putative RSE for testing as chimeric viruses, with the constructs carrying various combinations of 5′ and 3′ truncations. This was done in light of Wither's truncation studies of the RSV RSE that alluded to functional redundancy across the element; I wanted to see if the HIV putative RSE consists of redundant sub-elements across the 400 nt region as well.

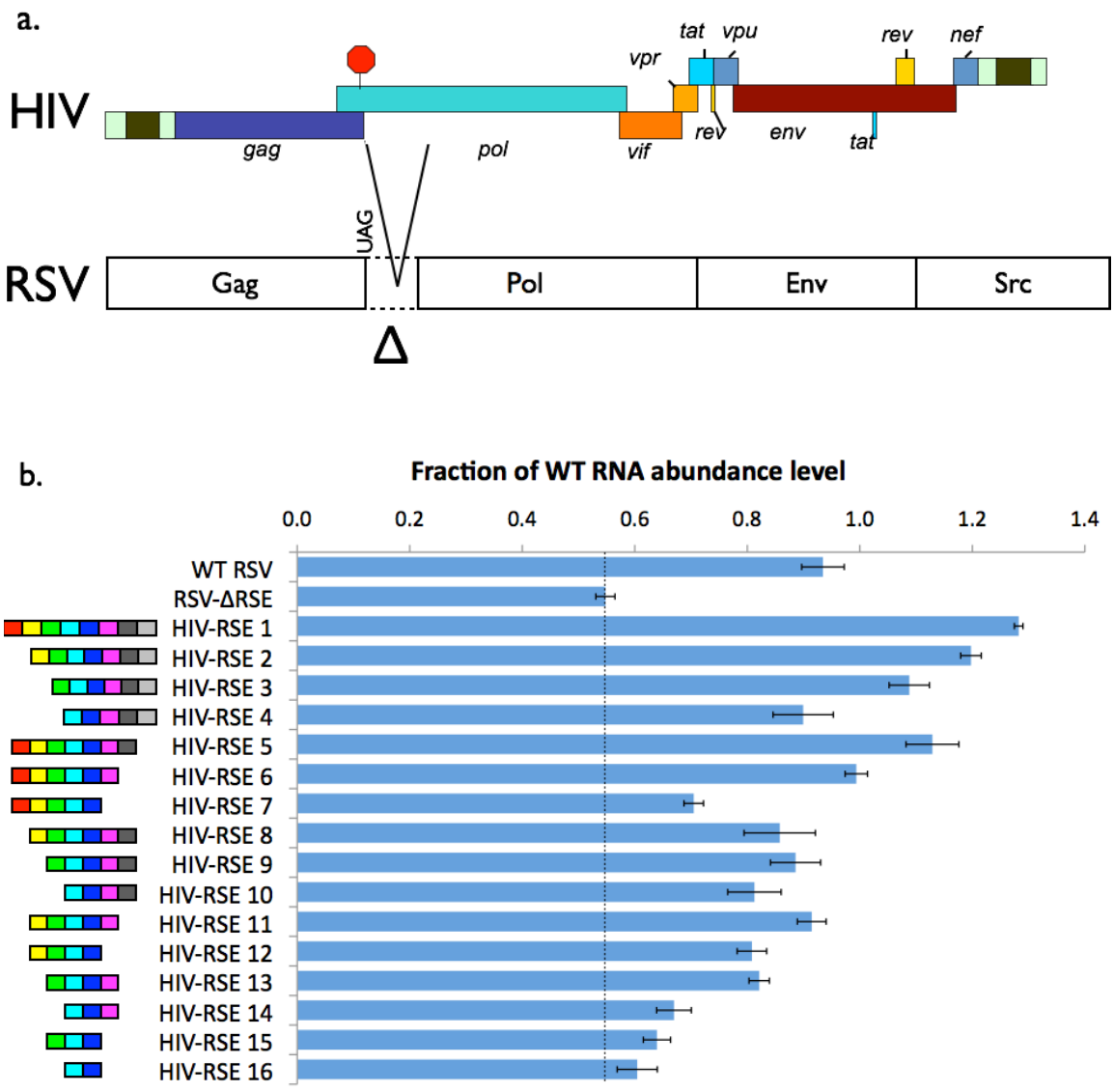


Figure 11: **HIV-1 3′ UTR confers RSE-like function in RSV. (a)** Schematic of HIV-RSV chimeric virus generation. **(b)** 400 nts of HIV-1 3′ UTR, as well as several partial fragments, are capable of stabilizing RSV RNA, in many cases performing better than the native RSV RSE. The data also suggests that there is functional redundancy across the HIV-1 putative RSE, similar to studies done on the RSV RSE. Each colored box left of the graph represents 50 nts of HIV 3′ UTR. Standard error represented on graph, n = 3

All sixteen variants of the HIV putative RSE were cloned into RSVΔRSE, and the stability of full-length RSV RNAs were evaluated by RPA. The results are shown in Figure 11b. WT HIV putative RSE appears to be able to stabilize full-length RSV RNA as a chimeric virus; in fact, the 400 nt HIV fragment appears to be able to stabilize the RSV RNA even better than the original WT RSV RSE. Similarly, the majority of the partial fragments of HIV putative RSE appear to confer partial to full functionality to RSV RNA in evading NMD; the protection against NMD is only lost with fragments that have been severely truncated and only comprise of 100-150 nts. This suggests that HIV may indeed have an RSE that is responsible for protecting its long 3′ UTR from being degraded by NMD. Previous work done in the lab also suggests that HIV putative RSE in the reverse orientation can protect long 3′ UTRs from NMD. [74] As such, I decided to move ahead with studying HIV in its natural context, by looking at full-length HIV RNA stability upon deletion of its putative RSE, as well as the effects of PTCs inserted into HIV *gag*.

With HIV, I left the first 6 nts after the *gag* stop codon intact as there were concerns that those nucleotides may be important for translation termination, and deleted the following 393 nts. I then tested HIV full-length RNA stability via RPAs.

HIV RNA expression level is extremely low in cells

I have had very limited success in assessing HIV RNA stability in cells due to its low expression levels. My initial hypothesis was that I was not getting efficient transfections of my plasmids into cells. In data not shown, I attempted to optimize transfections and RPAs of my HIV constructs into a number of different cell types (Jurkats, HeLas, HEK293s) using a variety of transfection methods (DEAE-Dextran, Lipofectamine 2000, Fugene 6, Fugene HD, electroporation) with varying DNA:transfection reagent ratios, but to little success. HeLas and HEK293s seem more amenable to transfections than Jurkats, and DEAE-Dextran yielded better results than the other transfection reagents and methods.

To rule out the possibility of inefficient transfection, I replaced the HIV control plasmid with human Globin and its corresponding Globin RPA probe that had previously been established in the lab. Figure 12a shows a sample RPA gel of this experiment. While Globin yields a very robust band, the HIV band is not present. This rules out transfection efficiency being the problem. At the same time, it is highly unlikely that the problem lies with the HIV RPA probe; excess undigested HIV probe is detectable, and since the probe hybridizes to its targets by simple base pairing, there should be no problems with hybridization kinetics. The only next logical conclusion is low expression.

Figure 12b shows a titration experiment, whereby I increased the amount of HIV DNA transfected up to 10 times higher than what is normally used, as well as increased the amount of transfection reagent used to see if I can increase transfection efficiency without compromising cell fitness. Increasing the amount of DNA transfected does

indeed increase RNA expression levels, and the extra DNA used did not seem to negatively affect the cells. Increasing the amount of DEAE-Dextran used, on the other hand, did yield stronger RPA bands, but the cells also looked progressively unhealthier with more DEAE-Dextran used. As such, I decided to continue my HIV experiments using 15µg each of the WT and the control HIV plasmids per 6 cm sample plate, and 10mg/mL DEAE-Dextran as normal.

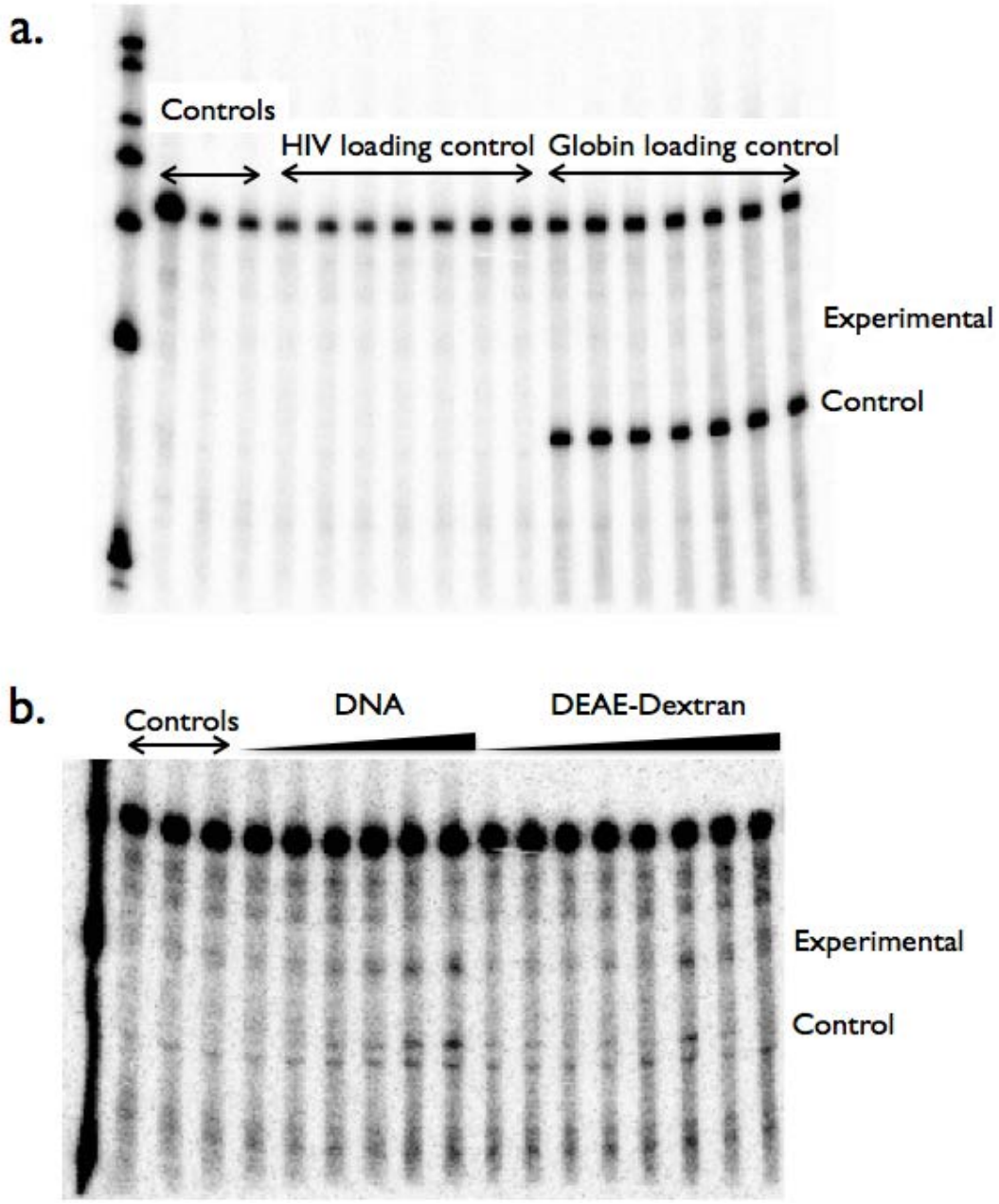


Figure 12: Low expression of HIV RNAs in HeLa cells. (a) HIV experimental and control RPA bands do not show up when transfected with the same amount of DNA as Globin. **(b)** HIV RNAs detectable at faint levels when DNA or Dextran amounts are drastically increased. Gel also shows HIV probe of the expected size with a strong signal, ruling out the possibility of a bad HIV probe.

HIV full-length RNA stability is not affected by the loss of its putative RSE or by PTCs inserted into gag.

Figure 13 shows the results of RPAs performed on HIVΔRSE in HeLa cells, along with a representative picture of an RPA gel. Steady-state abundance levels of HIV RNA with its putative RSE deleted is comparable to WT levels, showing that HIV may not have an RSE, or that it may not require the RSE for stability and protection against NMD. RNA levels did not appear to increase in the presence of DN-Upf1, thus showing that WT HIV full-length RNA itself is probably not a natural substrate of NMD that is already being rapidly degraded under steady-state conditions.

PTCs inserted into HIV gag were previously generated and tested by Weil, who observed that the PTCs do not affect the levels of unspliced HIV RNA in the cell. [74] There are, however, several caveats to this observation. Firstly, from the gel shown in Weil's thesis, it appears that he faced the same issue of low expression as I did: his HIV RPA bands were barely above background and I have hesitations as to how well they could be quantified. Secondly, the experiment was only performed once. It would be prudent to repeat the experiment a couple more times to ensure that the results can be replicated. Lastly, Weil did not delete the nucleocapsid from HIV to prevent RNA packaging. Instead, he merely harvested RNAs at an earlier timepoint (24 hrs post-transfection) in an attempt to examine the RNAs before they had a chance to be packaged. However, this means that the plasmids may only have had a short window to enter the nucleus and make RNAs. Furthermore, a portion of this already-limited RNA

pool may still have been successfully packaged into virions. This may result in the masking of any mutant phenotypes. As such, PTCs in HIV gag should be reexamined.

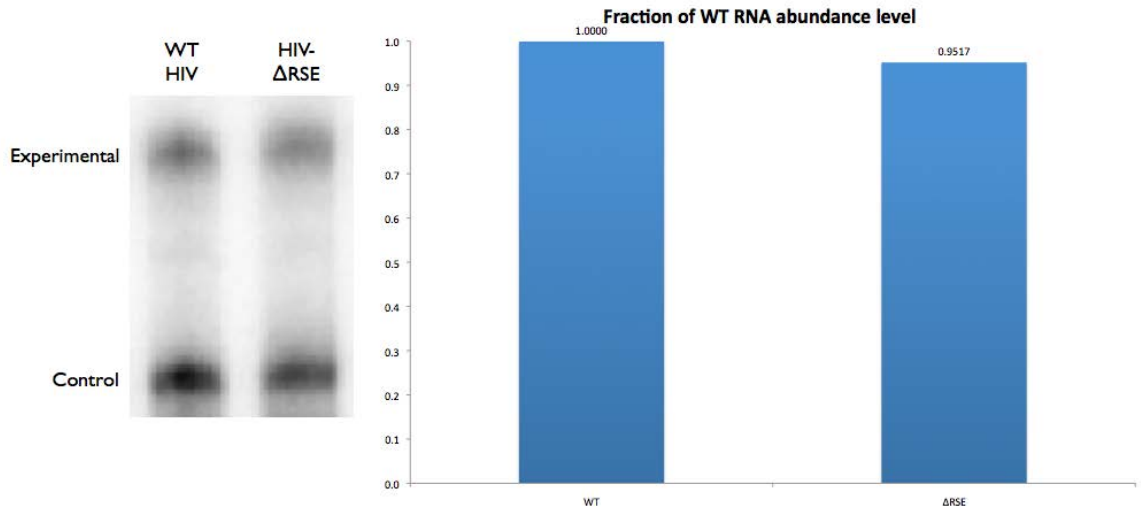


Figure 13: HIV may not have a putative RSE. HIVΔRSE RNA abundance levels are comparable to WT HIV.

Discussion

In light of these data, it appears that HIV does not require an RSE to protect its full-length RNA from degradation by NMD. There are a few possible explanations for our observations.

Firstly, the triggering of NMD is a translation-termination event; without translation, there will be no decay. To our knowledge, there has not been any studies done on how much HIV RNA is being translated. The ribosome footprinting experiment done in our lab has also been inconclusive. It therefore stands to reason that perhaps we do not see decay because the RNA is being translated at such a low frequency, and that NMD-mediated decay of the RNA happens at such a low rate, that RSE-mediated RNA stability is indistinguishable from WT stability levels.

Another possibility is that the HIV RNA has evolved another mode of protection against NMD that is distinct from RSE-mediated RNA stability, making the RNA completely immune to NMD regardless of the presence of PTCs and the absence of an RSE. If such a mechanism did exist, I would speculate that it evolved after the evolution of the RSE, since the HIV RSE does indeed have the capabilities of suppressing NMD of PTCs and long 3′ UTRs.

One last possibility, albeit quite unlikely, is that the HIV RNA itself does not bind Upf1, without which there will be no triggering of NMD. Even though the mechanism of how Upf1 is loaded onto RNAs is not well understood, Upf1 is a helicase that binds in a sequence-independent fashion and has been found to coat RNAs across the transcriptome. Therefore, it is hard to fathom that a specific RNA is able to prevent Upf1 binding

completely. Ajamian et al have shown that Upf1 is part of the HIV-1 RNP complex [75], although direct binding was not demonstrated. They also showed that Upf1 increases HIV-1 expression and translation independently from its role in NMD. The possibility exists that this increased expression may be an indirect consequence of NMD inhibition, such as in the case of HTLV-1 proteins inhibiting NMD to increase viral expression. [58, 59] However, global NMD is active in HIV-1 expressing cells, [75] so any increases in HIV expression due to NMD inhibition would likely be specific to the viral transcript.

Regardless, to continue studying HIV RNA stability, it is important that we overcome the technical challenge of low RNA expression levels, so that we are able to concretely observe and quantitate mutant phenotypes, if any. One possibility is to use qRT-PCRs in lieu of RPAs, as they are much more sensitive to small quantities of starting material. Alternatively, we can add a strong promoter in the constructs to promote higher rates of transcription.

400 nts of MMLV 3′ UTR immediately downstream of the gag stop codon can stabilize full-length RSV RNA in a chimeric construct.

Similar to our studies with HIV, we wanted to explore the presence of an MMLV putative RSE, by first testing the potential of the 3′ UTR immediately downstream of its *gag* stop codon in stabilizing full-length RSV RNA as a chimeric virus. This work had already been completed by Weil prior to me joining the lab, although the results were not reported. As shown in Figure 14, Weil inserted the 400 nt MMLV putative RSE in both the forward and reverse orientation behind a PTC in RSV *gag* and assessed steady-state full-length RSV RNA abundance via RPAs. MMLV putative RSE, when inserted in the forward orientation behind a PTC, is able to partially stabilize full-length RSV RNA, bringing the RNA abundance levels from approximately 25% for the PTC back up to 70%. When inserted in the reverse orientation however, the antisense MMLV putative RSE is no longer able to confer protection against NMD. This suggests that MMLV may indeed have a putative RSE that works in a similar fashion to the WT RSV RSE, and I decided to proceed with investigating the MMLV putative RSE in its natural context.

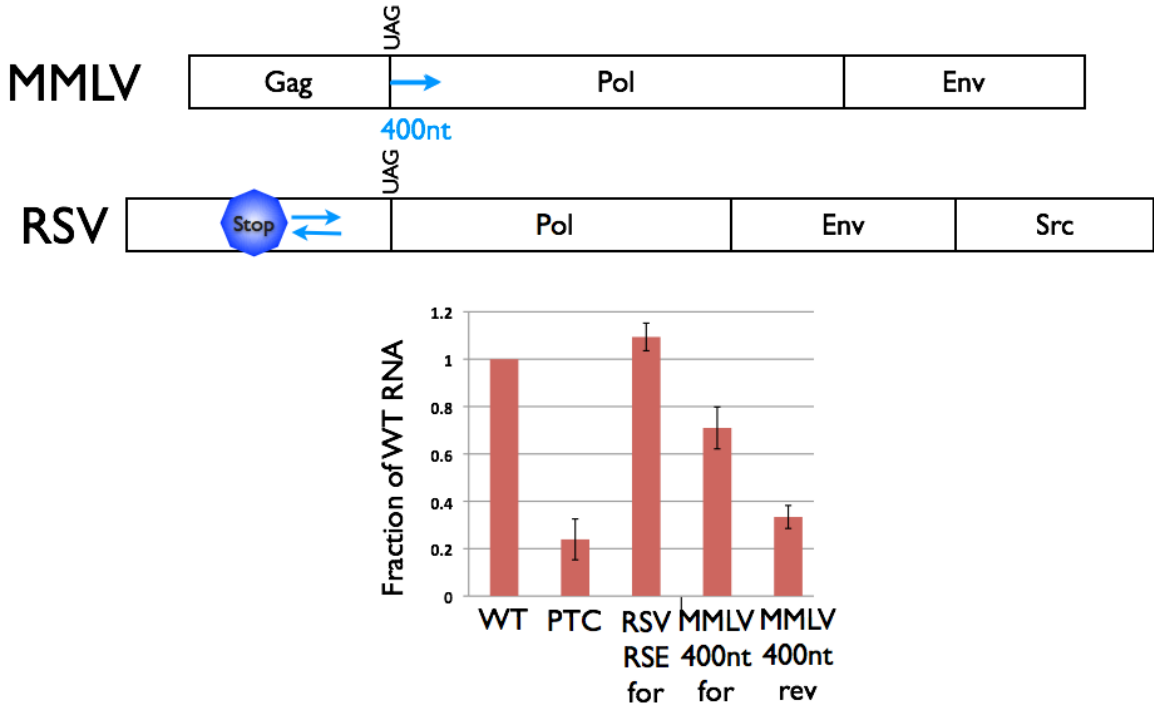


Figure 14: **MMLV 3′ UTR confers RSE-like function in RSV. Top:** Schematic of MMLV-RSV chimeric virus generation. 400 nts of MMLV 3′ UTR were inserted immediately downstream of a PTC in RSV, in both the forward and reverse orientation. **Bottom:** 400 nts of MMLV 3′ UTR in the forward orientation (for) are capable of stabilizing PTCs in RSV RNA, when inserted immediately downstream of the PTC. The reverse orientation (rev) of the MMLV 3′ UTR does not function in protecting the full-length RNA from NMD. Data from Weil.

MMLV RNAs that contain PTCs or lack its putative RSE are destabilized.

I generated a PTC in MMLV gag by mutating the C at nucleotide position 1326 to a T, thus creating an amber stop codon. This should place the PTC 645 nts upstream of the natural *gag* stop codon found at nts 1971–1973. However, since the downstream nucleocapsid zinc finger binding domain was deleted, the PTC is actually only 561 nts upstream of the gag stop codon. If the MMLV putative RSE is able to protect PTCs within 200 nts upstream of the gag stop codon as the RSV RSE does, the C1326T PTC should be located outside the range of protection, and the PTC-containing RNA should be significantly destabilized.

Figure 15a shows the RPA results of PTC-containing MMLV RNAs versus WT. Although the destabilization effect is not as dramatic as PTCs in RSV, PTC-containing MMLV RNAs do reproducibly show a statistically significant decrease in stability, with the RNA abundance at 79.7% that of WT levels. This suggests that MMLV full-length RNAs can be susceptible to degradation by NMD.

When the 400 nt putative RSE is deleted, the RNA destabilization phenotype is more severe, with the MMLVΔRSE steady-state RNA abundance levels falling to 60% of WT levels (see Figure 15b). This is in the range of what I commonly observe for RSVΔRSE and suggests that MMLV does indeed possess an RSE that, like the RSV RSE, is capable of protecting its long 3′ UTR from degradation by NMD.

However, attempts to confirm that NMD was responsible for the destabilization of MMLV RNAs containing PTCs or lacking the putative RSE have been inconclusive; co-

transfection with dominant negative Upf1 rescued the decay phenotype in 1 out of 3 trials (data not shown).

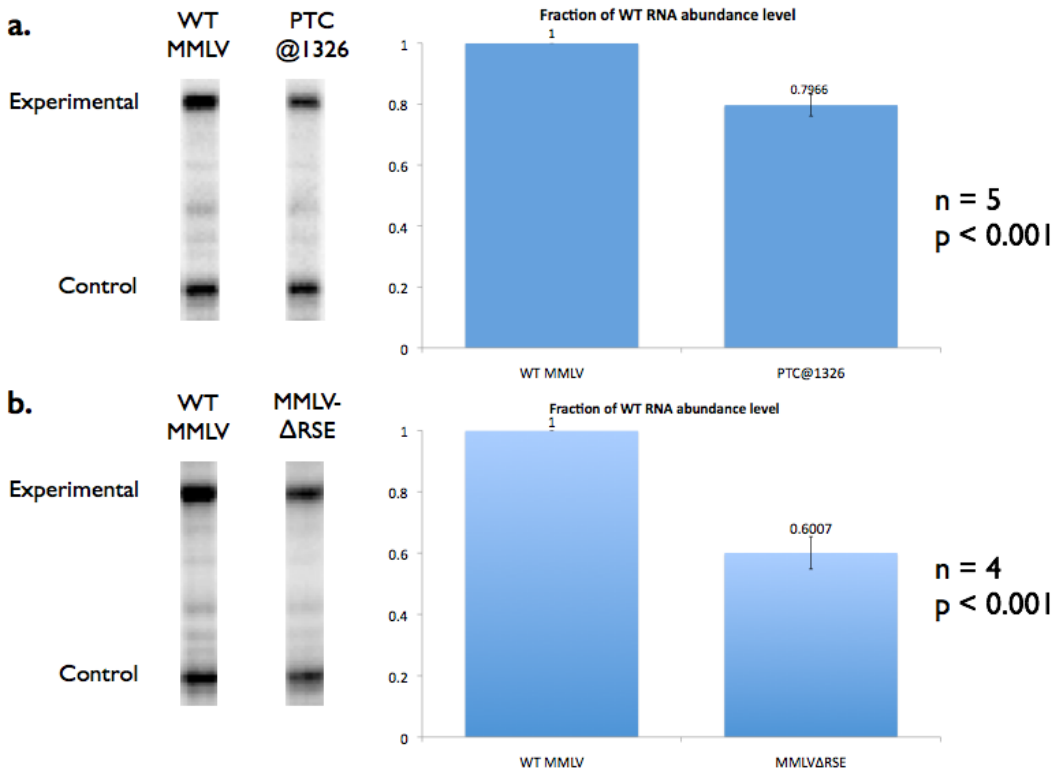


Figure 15: MMLV unspliced RNA is susceptible to NMD and may be protected by an RSE-like element. (a) A PTC generated at nt 1326 causes a mild but significant decrease of MMLV RNA levels, as assayed by RPA. Standard deviation represented on graph. n = 5; $p < 0.001$. **(b)** Deletion of the MMLV putative RSE causes MMLV RNA levels to decrease to 60% that of WT levels. n = 4; $p < 0.001$

Discussion

One explanation for why the PTC-containing MMLV RNA I tested was not degraded to a larger extent could be that the MMLV RSE is capable of providing a larger umbrella of protection as compared to the RSV RSE. Perhaps the MMLV RSE is simply capable of binding a larger number of proteins to prevent Upf1 binding over a larger region. Alternatively, the MMLV RSE may promote higher order protein complex formations that spread across large regions of the RNA. As such, PTCs at distances that are unprotected by the RSV RSE may still be fully or partially affected by the MMLV RSE.

In light of the discovery that PTBP1 binds RSV RSE and is important for the RSE's function in conferring protection against NMD, I wanted to see if PTBP1 might also be playing similar roles for HIV and MMLV. In Figure 16, I analyzed the binding site distribution across the HIV and MMLV genomes by adapting the python script that Nathan Lee previously wrote for analyzing the RSV genome. The green boxes outline the putative RSEs in the respective genomes, and it is clear that PTBP1 binding sites are not enriched within the putative RSEs nor in the general region downstream of the gag stop codons.

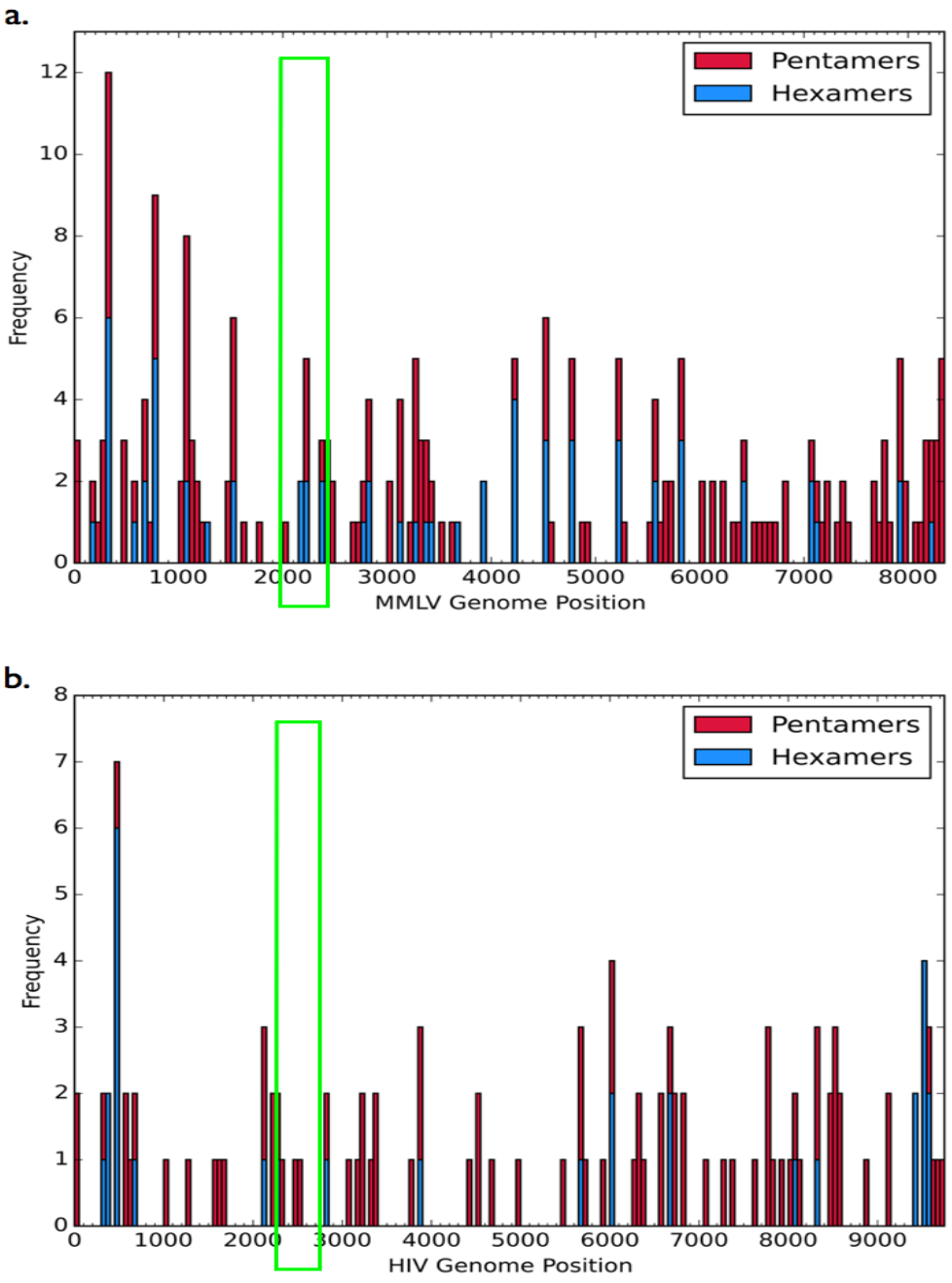


Figure 16: PTBP1 binding sites are not enriched downstream of the gag termination codon in both MMLV and HIV. Green boxes mark the respective putative RSE regions. **(a)** PTBP1 binding sites are evenly spread across MMLV genome. **(b)** PTBP1 binding sites are sparse in HIV genome.

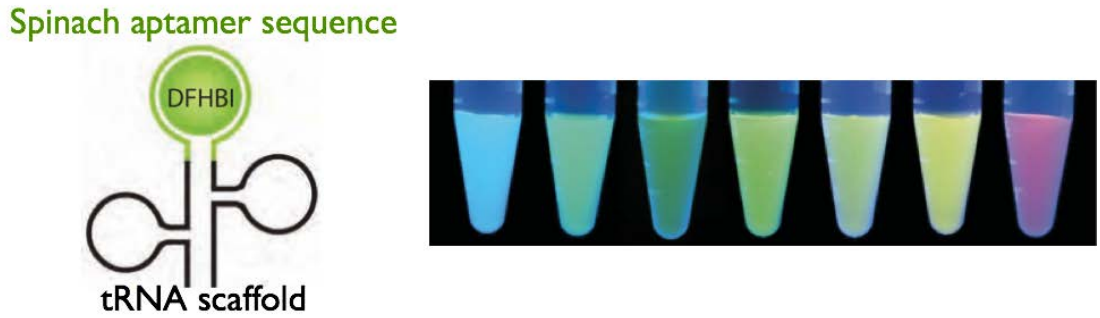
CHAPTER 4:

FUTURE DIRECTIONS

Developing a high-throughput assay to measure steady state RNA abundance levels

For our purposes, the RPA is a very robust and suitable technique for concurrently looking at steady-state RNA abundance levels across a limited number of samples. However, it is a long and labor-intensive protocol. As part of my project, I was interested in developing another assay that would allow me to screen mutants in a much more high-throughput fashion. For this, I turned to Spinach (see Figure 17).

Spinach is an RNA mimic of GFP, comprising of the Spinach RNA aptamer as part of a tRNA scaffold, and DFHBI, a drug that specifically binds the aptamer.[76] When DFHBI binds the 98 nt aptamer to form the Spinach complex, it fluoresces with excitation and emission wavelengths comparable to GFP, thus the output is detectable on microscopes and other platforms designed to measure GFP fluorescence. Additionally, DFHBI is non-toxic to cells and does not cause autofluorescence when unbound. This makes Spinach an ideal system for many live-cell applications, such as visualizing real-time RNA localization. Different combinations of RNA aptamers and drugs can yield a spectrum of fluorescence (Figure 17, right panel) that would enable tracking of multiple transcripts at the same time.



Adapted from Paige et al, Science, 2011

Figure 17: Spinach is an RNA mimic of GFP. Left: Spinach RNA aptamer in green, shown as part of a tRNA scaffold, binds DFHBI, forming a complex that behaves like GFP. Different combinations of aptamers and drugs yield a spectrum of florescence with different wavelengths and colors.

For my project, I was interested in using Spinach to directly assess transcript-specific steady-state RNA abundance levels in cells. The goal was to be able to grow and transfect cells in 96-well plates, and to measure RNA levels simply by measuring Spinach fluorescence via a qPCR machine. This would yield multiple benefits. Firstly, it will bring the actual assay time from 4 days for an RPA down to 30 minutes for a scan on the qPCR machine. In addition, I would be able to measure a much larger number of samples per round of experiment. It is also safer since we would bypass the process of generating, testing, and working with radioactive RNA probes. Lastly, it is more cost-effective, since the qPCR scan is much more sensitive than an RPA and we would be able to decrease the scale of our cell cultures, as well as save on the plethora of reagents needed to carry out the RPA.

I decided to tag my RSV RNAs with different numbers of Spinach concatamers and empirically decide the ideal number of concatamers to use. In the meantime, I transfected cells with the original expression vector obtained from the Jaffrey Lab. However, I only observed very faint fluorescence in a small number of cells under the microscope. Detection on the qPCR machine also yielded readings that were barely above background, and certainly insufficient to use as a reliable measure. Upon consultation with other labs trying to utilize Spinach during a conference, I found they were having similar issues. Spinach aptamers seem to only fold correctly 20% of the time, and the only labs that had marginal success in using Spinach had inserted over 20 contatamers into their transcript.

With that information at that time, I decided that it was no longer worthwhile to pursue Spinach for the time being and refocused on the RPA since it already works well. Nonetheless, I believe that Spinach is still a valuable tool that should be revisited and utilized in future if possible, especially now that the Jaffrey Lab has come up with updated versions of Spinach. These include the Spinach2, which folds better compared to its previous iteration[77], as well as Broccoli, another RNA mimic of GFP obtained by directed evolution via SELEX (Systematic Evolution of Ligands by Exponential Enrichment) followed by fluorescence- activated cell sorting (FACS)[78].

CHAPTER 5:

OVERALL DISCUSSION

Figure 18 shows our working model for general retroviral RSE-mediated RNA stability, in which long 3′ UTRs are prevented from triggering NMD by the occlusion of Upf1 molecules binding near the natural gag stop codon. For RSV, the protein in question binding the RSE is PTBP1; however, it is probably not the protein binding the MMLV RSE, or the HIV RSE if it exists. There is no reason to suspect that the same protein has to be involved across all retroviruses either; any competing factors that can bind the region to occlude Upf1 binding would be able to fulfill the same role.

It seems reasonable to suspect that this is how PABP stabilizes mRNAs without long 3′ UTRs: PABP occludes Upf1 from binding near natural termination codons, thus providing contextual clues to prevent decay of normal transcripts while still allowing PTCs to get detected and degraded via NMD. This explains why RNA stabilization is achieved in PTCs and in RNAs with long 3′ UTRs when PABP is artificially brought closer to the termination codon in question.

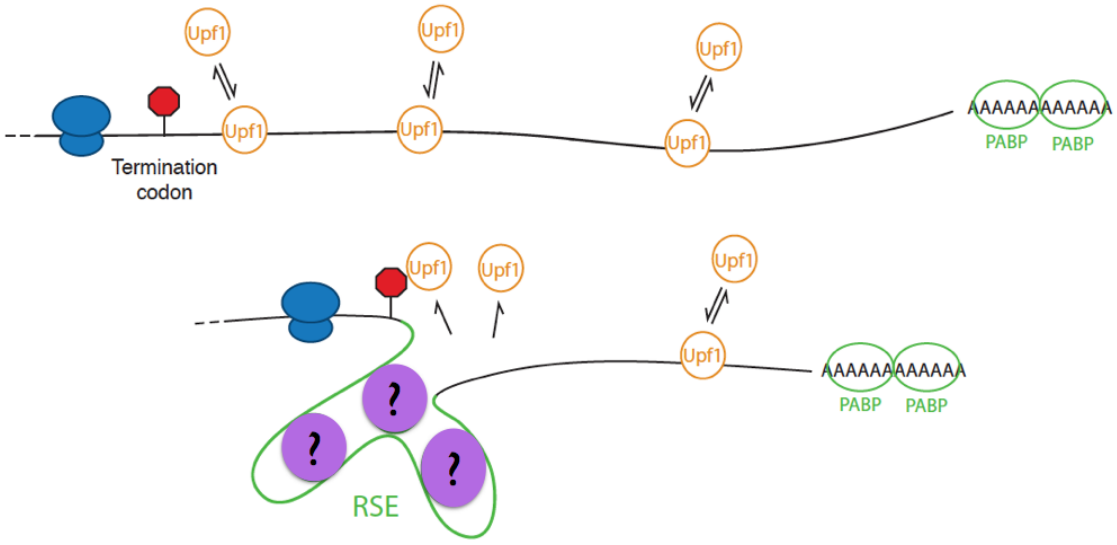


Figure 18: **Retroviral RSEs protect their RNAs from NMD by inhibiting Upf1 binding.** Different RSEs may bind different proteins to establish higher-order mRNP structures in the vicinity of the stop codon. This prevents Upf1 from binding near the termination codon, and disrupts the Upf1-ribosomal interaction needed to trigger NMD. The NMD pathway thus judges the 3′ UTR to be short and fails to degrade the mRNA.

Another piece of evidence in support of this hypothesis, albeit weaker, is that heterogeneous nuclear ribonucleoprotein L (hnRNPL) appears to be able to succeed PTBP1's capacity in RSE-mediated RNA stability. In data not shown, a previous iteration of RSVΔPTB mutant that we previously generated exhibited WT RNA stability. The mutant was created by mutating Us to As; in doing so, we inadvertently created CA-rich sequences, which are hnRNPL binding sites. Confounding this analysis, however, is that hnRNPL has previously been shown to contribute to RNA stability, and it can also interact with PTBP1. As such, it is unclear whether the RNA stabilization effects seen in the RSEΔPTB+hnRNPL mutant was a result of RSE conferring RNA stability, hnRNPL's inherent capabilities in stabilizing RNAs, or if hnRNPL is recruiting PTBP1 to the site.

Nonetheless, the 3′ UTR region in the retroviruses studied for this project each code for different proteins (reverse transcriptase for RSV, and different regions of protease for HIV and MMLV). Thus, they have different nucleotide sequences and RNA 2º structures, so we would expect their respective RSEs to bind different proteins. As in the discussion in Chapter 3, there is no reason to suspect that only one protein is involved in each of these RSEs. For instance, with regards to the RSV RSE, PTBP1 could be working in concert with serine/arginine-rich splicing factors, which were also candidate protein interactors in my mutagenesis screen and our collaborators' mass spectrometry screen, to coat the RNA near the termination codon. Alternatively, the RSE could recruit factors to increase translation termination efficiency and speed up ribosome disassembly to prevent decay complex formation. Thus, the sequence immediately downstream of the correct viral termination codon, and its associated ribonucleoproteins, is important in

validating that termination codon and preventing NMD. Different protein interactions and complex formations may affect the range of protection against NMD that each RSE can provide.

Furthermore, as mentioned in Chapter 1, the 400 nt region is an arbitrary designation assigned when the RSE project was first started, based on restriction sites [61]. At that time, we only knew that we were looking at an RNA element downstream of the RSV gag stop codon that worked in a distance-dependent fashion, with no knowledge of its size, structure, or mechanism.

Going forward, I think it is important to reconsider our understanding of RSEs and refine our approach when studying them. For example, the RSE for a particular retrovirus may be bigger or smaller than 400 nts. Instead of looking purely at a particular size, we might be better served if we looked for enrichment of protein and protein complexes in the regions surrounding the termination codon. Another approach would be to consider RNA structure when outlining the region of interest. For instance, according to SHAPE data from the Weeks lab [79, 80], the 400 nts currently designated as the HIV RSE represents 1.5 stemloops; it may be more appropriate to consider a 500-550 nt region so that we are studying two complete stemloops. Our structural RNA collaborators are also working on HIV RSE RNA, so future refinements based on their data would also be beneficial.

One other thing to note is that the retroviral mechanism of protecting long 3′ UTRs from NMD may be conserved in a subset of human genes. Our collaborators in the Hogg Lab performed a global analysis of the role of PTBP1 in regulating RNA stability.

As shown in Figure 19a, knocking down PTBP1 in human cells by siRNAs causes a decrease in RNA abundance in 535 genes, denoted by the red dots. When a double knockdown of PTBP1 and Upf1 was performed, the RNA abundance levels were rescued in 188 of these genes, as represented in green. Figure 19b is a continuous distribution function plot of 3′UTR lengths. Compared to all genes, these 188 genes on average have longer 3′UTRs. Put together, this tell us that these 188 genes have long 3′UTRs that are protected against Upf1-mediated NMD by PTBP1, which suggests that they may have RSE-like elements that function similar to the RSV RSE.

In the same vein, a very recent study from the Lykke-Anderson lab has shown that for a subset of human mRNAs with long 3′ UTRs, there is a cis-acting element within 200 nts of the natural termination codon that is important for RNA stability. [68] The 200 nt region tends to be AU-rich for 3′ UTRs that were inhibitory to NMD, although they have not yet identified specific proteins involved.

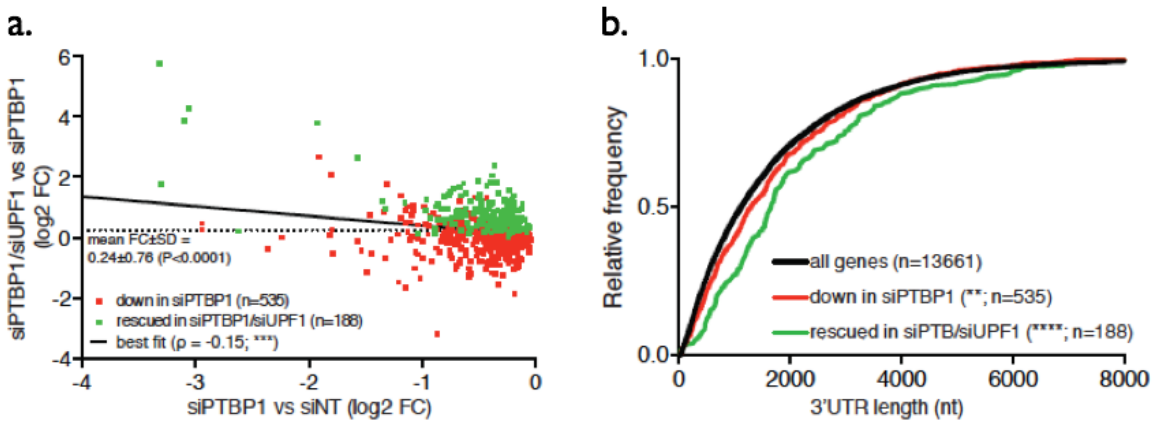


Figure 19: PTBP1 protects many human mRNAs with long 3′UTRs from NMD. (a). RNAseq identifies mRNAs protected by PTBP1. For 535 mRNAs reproducibly down-regulated by PTBP1 depletion, fold change comparing siPTBP1/siUPF1 to siPTBP1 alone was plotted vs. fold change comparing siPTBP1 to siNT. mRNAs reproducibly rescued in siPTBP1/siUPF1 co-depletion are shown in green. Dashed line indicates mean fold change (P<0.0001 in K-S test comparing to all genes); solid line indicates best fit determined by least squares method (Spearman's ρ = -0.15; ***: P<0.001). **(b)** mRNAs protected from NMD by PTBP1 have long 3′UTRs. Continuous distribution function plot of annotated 3′UTR lengths among all mRNAs, mRNAs down-regulated by PTBP1 depletion, and mRNAs rescued by co-depletion of PTBP1 and UPF1 in RNA-seq analyses. Statistical significance was evaluated by K-S test, comparing the indicated mRNA sets to the distribution of 3′UTR lengths among all mRNAs (**: P<0.01; ****: P<0.0001). Data from Ge and Hogg.

With increasing gene complexity, organisms have had to evolve mechanisms for them to fine tune gene expression, such as increasing 3′ UTR lengths in their RNAs to accommodate more regulatory elements. The RSE is thus an ideal tool for these organisms to exploit, as it stabilizes natural termination codons in RNAs with long 3′ UTRs while still allowing premature termination codons to be degraded. We think it is likely that this mode of protection against NMD is conserved in other organisms beyond humans and retroviruses.

We also believe that the RSE has many important potential biotechnical and gene therapy applications. The most straightforward application of the RSE would be in RNA therapeutics, whereby RNAs are delivered directly to cells in order to generate proteins of interest. The RSE can be used to stabilize these RNAs, should they require long 3′ UTRs to achieve better regulation. In addition, RSE-mediated RNA stability is an easy post-transcriptional method to increase protein yield for large-scale manufacturing and purification purposes. For gene therapy purposes, being able to deliver multiple genes on a single DNA construct is preferable to delivering different genes in trans, as it is more costly and technically challenging to achieve efficient transfection rates across a large number of cells. The RSE would then be a crucial element to include in the construct after each ORF, so that the RNA does not undergo decay.

As such, the RSE is an important element that should continue to be studied in greater detail for us to fully understand the mechanisms of its function. This will undoubtedly aid us in our understanding of NMD, and provide a useful biomedical tool in the quest to cure human diseases.

CHAPTER 6:

MATERIALS AND METHODS

Cell cultures and transfections

Cell cultures and transfections for RSV

Secondary chicken embryo fibroblasts (CEFs) were grown at 39°C under 5% CO_2 in 211 medium (199 medium supplemented with 2% tryptose phosphate broth, 1% calf serum, 1% chick serum, and 1% antimicrobial-antimycotic). Transient transfections were performed using DEAE-dextran at 100-200 µg/mL (varied based on manufacturer due to differential toxicity effects on cells) in serum-free 199 medium as previously described.[81] Cells were plated in 6 cm dishes and transfected with 1.5 µg of each construct used (3-4.5µg total DNA per plate) at 60% confluency. Total cellular RNA was harvested 40-48 hrs post-transfection from CEFs using RNA-Bee (Tel-test) as per manufacturer's recommendations.

Cell cultures and transfections for HIV

HeLa and HEK293 cells were grown at 37°C under 5% CO_2 in DMEM supplemented with 10% fetal bovine serum and 1% antimicrobial-antimycotic. Transient transfections were performed as described above for CEFs, using 15 µg DNA per 6 cm cell culture plate. Transfections using DEAE-dextran were performed in serum-free DMEM. For transfections using Lipofectamine 2000, Fugene 6, or Fugene HD, Opti-MEM was used for DNA-transfection reagent complex formation, before the complexes were added directly to cells in growth medium as per manufacturer's recommendations.

Jurkat cells were grown at 37°C under 5% CO_2 in RPMI-1640 medium supplemented with 10% fetal bovine serum and 1% antimicrobial-antimycotic. Transient transfections were performed using Fugene 6 or Fugene HD as described above for HeLas/HEK293s.

Cell cultures and transfections for MMLV

NIH3T3 cells were grown at 37°C under 5% CO_2 in DMEM supplemented with 10% calf serum and 1% antimicrobial-antimycotic. Transient transfections were performed in serum-free DMEM as described above for CEFs.

Viral constructs

RSE mutagenesis screen

All RSV nts correspond to NCBI reference sequence: NC_001407.1.

The 10.8 plasmid, ΔRSE plasmid, and 1547C construct previously described[22, 65, 82] were used as RSV WT, RSVΔRSE and RSV control plasmids in this study.

The previously described RSV WT E/S vector[66], which contains an EagI restriction site at nt position 2492-2497 and an SpeI restriction site at nt position 2886-2891, was used to generate minRSE mutants.

RSV WT E/S version 2 vector, a variant of the RSV WT E/S vector whereby the EagI restriction site was generated at nt position 2572-2577, was previously made by Withers (unpublished). This vector was used to generate C-frag mutants.

Mutants were generated by mutagenic PCR as previously described[83]. Combine 10 µL of 10x mutagenic PCR buffer (70 mM $MgC1_2$, 500 mM KC1, 100 mM Tris (pH 8.3 at 25ºC), 0.1% (wt/vol) gelatin), 10 µL of 10x dNTP mix (2 mM dGTP, 2 mM dATP, 10 mM dCTP, 10 mM TTP), 30 pmoles of each primer, 20 fmoles of template DNA, and bring to 88 µL with dH_2O. Mix well. Add 10 µL of 5 mM $MnCI_2$ and mix well. Add 2 µL (5 U) Taq polymerase for a final volume of 100 µL. Incubate for 30 cycles of 94ºC for 1 min, 45ºC for 1 min, and 72ºC for 1 min, without a "hot start" procedure or a prolonged extension time at the end of the last cycle. DNA template for the mutagenesis was first amplified from the 10.8 RSV viral plasmid and purified from an agarose gel using the Fermentas gel extraction kit. Mutagenic PCR primers, designed to include an EagI restriction site in the forward primer and an Spel restriction site in the reverse primer, were then used to generate amplicons through 3 rounds of mutagenic PCRs. Amplicons were cloned into their respective RSV vectors as described below.

Cloning amplicons into RSV vectors

Amplicons and WT vectors were digested with EagI and Spel, and the vectors were treated with calf intestinal phosphatase. Digested amplicons and vectors were gel-purified and ligated before being transformed into *E. coli*. Clones were screened by colony PCR or grown for plasmid purification followed by restriction digests. Selected clones were then confirmed by sequencing.

RSE-PTB variants in RSV constructs

Plasmids containing RSEΔPTB, RSEΔPTB+3xPTB, and RSEΔPTB+6xPTB were obtained from Hogg and Ge. They abolished putative PTBP1 binding sites in the RSV

RSE by mutating C to A and U to G to create RSEΔPTB. 3 and 6 previously verified exogenous PTPBP1 binding sites were added back to RSEΔPTB to form RSEΔPTB+3xPTB and RSEΔPTB+6xPTB respectively.

RSE variants were amplified from the obtained plasmids by PCR, and the amplicons were cloned into RSV WT E/S vector as described above.

HIV-RSV chimeric virus constructs

All HIV nts correspond to HIV-1/HTLV-III/LAV reference genome HIV-1 HXB2 strain, NCBI reference sequence: GenBank: K03455.1.

Plasmid pNL43-ΔE-GFP was a gift from Robert Siliciano that contains a copy of HIV-1 proviral DNA on a pUC18 backbone, with EGFP inserted to split and inactivate the viral envelope.

HIV-RSV chimeric virus constructs were generated by amplifying HIV 3′ UTR from pNL43-ΔE-GFP. PCR primers were designed to include an EagI restriction site in the forward primer and an Spel restriction site in the reverse primer. Amplicons were cloned into RSV WT E/S vector as described above.

HIV constructs

To generate HIV WT vector, I deleted a 105 nt region (nts 1963-2067) in gag containing the nucleocapsid (NC) zinc finger binding domains from pNL43-ΔE-GFP. 5′ phosphates were added to PCR primers using T4 polynucleotide kinase. Primers were purified by phenol-chloroform extraction and used for inverse PCR. The PCR product was digested

with DpnI and ligated before being transformed into *E. coli*. Clones were screened by colony PCR or grown for plasmid purification followed by restriction digests. Selected clones were then confirmed by sequencing. This HIVΔNC_{CCHC} plasmid is used as the HIV WT vector for my project.

A control plasmid was generated by deleting a larger, 147 nt region (nt 1921-2067) from pNL43-ΔE-GFP gag, also spanning the NC zinc finger binding domains. This was done using inverse PCR as described above for HIV WT vector, and the resulting HIVΔNC_{CCHC+} plasmid is used as the HIV control vector for my project.

MMLV constructs

All MMLV nts correspond to NCBI reference sequence: NC_001501.1.

Plasmid pNCS was a gift from Stephen Goff that contains an infectious copy of MMLV proviral DNA on a pBR322 backbone modified to include an SV40 origin. I deleted an 84 nt region (nts 1836-1919) in gag containing the NC zinc finger binding domain from pNCS, using inverse PCR as described above for HIV WT vector. This MMLVΔNC_{CCHC} plasmid is used as the MMLV WT vector for my project.

A control plasmid was generated by deleting a 171 nt region (nt 3072-3242) in pol. The MMLV WT vector was digested with AvrII. The larger band was gel-purified before being transformed into *E. coli*. Clones were screened by colony PCR or grown for plasmid purification followed by restriction digests. Selected clones were then confirmed by sequencing. This MMLVΔNC_{CCHC}Δ3072-3242pol plasmid is used as the MMLV control vector for my project.

Upf1 constructs

WT Upf1 (hUpfl) and dominant negative Upf1 (RR857GA) constructs were gifts from Hal Dietz, Johns Hopkins University School of Medicine, and are described previously [84].

RNase protection assays (RPAs)

RSV RPAs were performed using an RSV *gag* riboprobe previously described[64] that yields a 627 nt band when undigested, 430 nt band for WT RSV, and a 387 nt band for RSV control.

HIV RPAs were performed using an HIV *gag* probe. The 332 nt probe is complementary to nt 1736–2067 of pNL43-ΔE-GFP HIV RNA and spans the NC deletion. WT HIV yields a 227 nt protected product, and HIV control yields a 185 nt protected product.

MMLV RPAs were performed using an MMLV *pol* probe. The 228 nt probe is complementary to a 197 nt region spanning nt 2915–3111 of pNCS MMLV RNA. WT MMLV yields a 197 nt protected product, and MMLV control yields a 164 nt protected product.

In vitro transcription of antisense viral probes were performed from a T7 DNA template and radiolabeled with [α-^{32}P]GTP as previously described[64].

Total cellular RNA (10 μg) was resuspended in 30 μL of 80% formamide hybridization solution (80% vol/vol deionized formamide, 40 mM piperazine-N,N'-bis(2-

ethanesulfonic acid) pH 6.7, 1 mM EDTA, 0.4 mM NaCl) containing ~250 000 cpm of probe. RNAs were denatured at 95°C and incubated at 42°C for 16 hrs. 300 µL of RNase digestion buffer (10 mM Tris-HCl pH 7.5, 300 mM NaCl, 5 mM EDTA, 10 U/mL RNase Tl, and 5 µg/mL RNase A) was added and then incubated at 33°C for 45 min. Sodium dodecyl sulfate (SDS) and proteinase K were added to final concentrations of 0.6% vol/vol and 0.14 mg/mL, respectively, followed by a 20 min incubation at 37°C to stop the RNase digestion. The samples were extracted with an equal volume of phenol-chloroform-isoamyl alcohol (25:24:1) followed by ethanol precipitation. RNAs were resuspended in 95% formamide loading dye (95% vol/vol deionized formamide, 0.02% bromophenol blue, 0.02% xylene cyanol) and denatured for 3 min at 95°C. Samples were electrophoresed on a 6% acrylamide-8M urea sequencing gel. RNA levels were quantified using a Typhoon Phosphoimager and GE Imagequant software.

Electrophoretic mobility shift assay (EMSA)

Sense viral riboprobes were generated via in vitro transcription from a T7 DNA template and radiolabeled with [α-^{32}P]GTP as described for RPA probe generation.

EMSAs were performed as previously described[85]. 1 µL of 38.4 µM PTBP1 protein was added to 9 µL autoclaved binding buffer stock (20 mM HEPES pH 7.6, 80 mM glutamic acid, 80 mM KOH, 6 mM $MgCl_2$, 20% glycerol) supplemented with 1 mM DTT and 0.02% SDS, and incubated at 30ºC for 8 min. 1 µL of 40 ng/µL RSE variant RNA was added to the solution and incubated at 30ºC for 15 min. The samples were then

incubated on ice for 5 min. 2 μL loading dye (0.25% bromophenol blue, 89 mM Tris pH 7.6, 89 mM boric acid, 2 mM EDTA, 10% glycerol) was added and the samples were electrophoresed on a 4% acrylamide gel (pre-ran for ≥30 mins) for 6 hr at 150V at 4ºC.

Tagging RSV RNAs with Spinach aptamers

pAV-5S-Spinach plasmid was a gift from Samie Jaffrey, Cornell University. The 171bp Spinach aptamer-tRNA scaffold was amplified by PCR using kinased primers, and circularized by ligation. Concatamers were generated via rolling circle amplification (GE illustra TempliPhi 100 Amplification kit) and re-amplified with primers containing restriction sites. Spinach aptamer-tRNA scaffold concatamers were cloned into pol by restriction digest and ligation into the RSV WT plasmid. Clones were screened by colony PCR or grown for plasmid purification followed by restriction digests. Selected clones were then confirmed by sequencing.

APPENDIX I:

PRIMERS LIST

RSV

Primer	Sequence
C-frag mutagenesis F	GATTGACCAGCGGCCG
C-frag mutagenesis R	CCTGGTTATTCACAGAGACTAGT
minRSE mutagenesis F	CGACGGCCGGACCAGTGGCCC
minRSE mutagenesis R	CGCACTAGTGGAACAAGCTTGGCGTTA
RSEΔPTB & +3xPTB F	CAGGGCGGCCGTGATCAAGGTTGCGCTACATC
RSEΔPTB+6xPTB F	TAGGGCGGCCGTGATCAAGGTTGCGCTACATTC
RSEΔPTB variants R	TGGGGACTAGTGTAAATGCAAAAGCTTCGCGAT C
RSV RPA gag probe F	CGATCGATAGCTAGCATCGACTACGACTGCGCG GCGTCGGCTCTC
RSV RPA gag probe T7 R	TAATACGACTCACTATAGGGTCATCGCTAACGA GACGGC
EMSA RSE RNA probe T7 F	TAATACGACTCACTATAGGGGGAGGGCCACTGT TCTC
EMSA RSE RNA probe R	GTAAATGCAAAAGCTTCGCG
Spinach-tRNA 5′ stem F	GCCCGGATAGCTCAGTCG
Spinach-tRNA 3′ stem R	TGGCGCCCGAACAGG

HIV

Primer	Sequence
2303 EagI F HIV-RSV	GACGACCGGCCGGCAACTAAAGGAAGCTCTATTAGATACAG
2344 EagI F HIV-RSV	GACGACCGGCCGCAGTATTAGAAGAAATGAGTTTGCCAG
2397 EagI F HIV-RSV	GACGACCGGCCGGGAATTGGAGGTTTTATCAAAGTAAGAC
2441 EagI F HIV-RSV	GACGACCGGCCGCATAGAAATCTGTGGACATAAAGCTATAGG
2539 SpeI R HIV-RSV	ATGATGACTAGTGTGCAACCAATCTGAGTCAAC
2592 SpeI R HIV-RSV	ATGATGACTAGTCTGGCTTTAATTTTACTGGTACAGTC
2634 SpeI R HIV-RSV	ATGATGACTAGTCTTCTGTCAATGGCCATTGTTTAAC
2692 SpeI R HIV-RSV	ATGATGACTAGTGAAATTTTCCCTTCCTTTTCCATCTC
HIVΔ$NC_{CCHC/CCHC+}$ F	ACTGAGAGACAGGCTAATTTTTTAGG
HIVΔNC_{CCHC} R	CTTAACAGTCTTTCTTTGGTTCCTAAAATTG
HIVΔNC_{CCHC+} R	CATTATGGTAGCTGGATTTGTTACTTG
HIVΔRSE F	CAAAAATTGGGCCTGAAAATC
HIVΔRSE R	CTATCTTTATTGTGACGAGGG
Sequencing HIVΔRSE	CTTCAGAGCAGACCAGAG
HIV RPA gag probe F	GGATGACAGAAACCTTGTTGGTC
HIV RPA gag probe R	TAATACGACTCACTATAGGGACAATCTTTCATTTGGTGTCCTTCC

MMLV

Primer	Sequence
MMLVΔNC$_{CCHC}$ F	GGACCAAGACCCCAGACC
MMLVΔNC$_{CCHC}$ R	TTCTCCTCCCTGTCTATCCTG
PTC C1326T F	CCAGACTGGGATTACACCACCTAGGCAGGTAG GAACCACC
PTC C1326T R	GGTGGTTCCTACCTGCCTAGGTGGTGTAATCCC AGTCTGG
MLVΔRSE F	TCAAAAGAGCCAGATGTTTCTC
MLVΔRSE R	CTAGTCATCTAGGGTCAGGAG
MMLV RPA pol probe F	CGATCGATAGCTAGCATCGACTACGACTGCGCA GTCCCACCCTGTTTGATG
MMLV RPA pol probe T7 R	TAATACGACTCACTATAGGGCTTTCTTGGCCGA GG

APPENDIX II:

PERMISSION LETTERS

ELSEVIER LICENSE
TERMS AND CONDITIONS

May 28, 2015

This is a License Agreement between Bao Lin Quek ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Bao Lin Quek
Customer address	4000 N Charles St Apt 1206 BALTIMORE, MD 21218
License number	3637940830366
License date	May 28, 2015
Licensed content publisher	Elsevier
Licensed content publication	Current Opinion in Microbiology
Licensed content title	Retroviral strategy to stabilize viral RNA
Licensed content author	None
Licensed content date	April 2014
Licensed content volume number	18
Licensed content issue number	n/a
Number of pages	5
Start Page	78
End Page	82
Type of Use	reuse in a thesis/dissertation
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Title of your thesis/dissertation	Characterization of Retroviral RNA Stability Elements

Expected completion date	Jun 2015
Estimated size (number of pages)	80
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD
VAT/Local Sales Tax	0.00 USD / 0.00 GBP
Total	0.00 USD
Terms and Conditions	

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment

NATURE PUBLISHING GROUP LICENSE TERMS AND CONDITIONS

May 28, 2015

This is a License Agreement between Bao Lin Quek ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3637941041908
License date	May 28, 2015
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature Reviews Microbiology
Licensed content title	Studies of endogenous retroviruses reveal a continuing evolutionary saga
Licensed content author	Jonathan P. Stoye
Licensed content date	May 8, 2012
Volume number	10
Issue number	6
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
High-res required	no
Figures	Figure 1
Author of this NPG article	no
Your reference number	None
Title of your thesis / dissertation	Characterization of Retroviral RNA Stability Elements
Expected completion date	Jun 2015
Estimated size (number of pages)	80
Total	0.00 USD
Terms and Conditions	

Terms and Conditions for Permissions

THE AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE LICENSE TERMS AND CONDITIONS

Jul 22, 2015

This is a License Agreement between Bao Lin Quek ("You") and The American Association for the Advancement of Science ("The American Association for the Advancement of Science") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by The American Association for the Advancement of Science, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3674261217190
License date	Jul 22, 2015
Licensed content publisher	The American Association for the Advancement of Science
Licensed content publication	Science
Licensed content title	RNA Mimics of Green Fluorescent Protein
Licensed content author	Jeremy S. Paige, Karen Y. Wu, Samie R. Jaffrey
Licensed content date	Jul 29, 2011
Volume number	333
Issue number	6042
Type of Use	Thesis / Dissertation
Requestor type	Scientist/individual at a research institution
Format	Print and electronic
Portion	Figure
Number of figures/tables	1
Order reference number	None
Title of your thesis / dissertation	Characterization of retroviral RNA stability elements
Expected completion date	Jul 2015
Estimated size(pages)	126
Total	0.00 USD
Terms and Conditions	

American Association for the Advancement of Science TERMS AND CONDITIONS

Regarding your request, we are pleased to grant you non-exclusive, non-transferable permission, to republish the AAAS material identified above in your work identified above, subject to the terms and conditions herein. We must be contacted for permission for any uses other than those specifically identified in your request above.

The following credit line must be printed along with the AAAS material: "From [Full

REFERENCES

1. Stoye JP: **Studies of endogenous retroviruses reveal a continuing evolutionary saga.** *Nat Rev Microbiol* 2012, **10**:395–406.

2. Coffin JM, Hughes SH, Varmus HE: *Retroviruses*. Cold Spring Harbor Laboratory Press; 1997.

3. Hatziioannou T, Goff SP: **Infection of nondividing cells by Rous sarcoma virus.** *J Virol* 2001, **75**:9526–31.

4. Yamashita M, Emerman M: **Retroviral infection of non-dividing cells: old and new perspectives.** *Virology* 2006, **344**:88–93.

5. Werner S, Hindmarsh P, Napirei M, Vogel-Bachmayr K, Wöhrl BM: **Subcellular localization and integration activities of rous sarcoma virus reverse transcriptase.** *J Virol* 2002, **76**:6205–12.

6. Scheifele LZ, Garbitt RA, Rhoads JD, Parent LJ: **Nuclear entry and CRM1-dependent nuclear export of the Rous sarcoma virus Gag polyprotein.** *Proc Natl Acad Sci U S A* 2002, **99**:3944–9.

7. Garbitt-Hirst R, Kenney SP, Parent LJ: **Genetic evidence for a connection between Rous sarcoma virus gag nuclear trafficking and genomic RNA packaging.** *J Virol* 2009, **83**:6790–7.

8. Frischmeyer PA, van Hoof A, O'Donnell K, Guerrerio AL, Parker R, Dietz HC: **An mRNA surveillance mechanism that eliminates transcripts lacking termination codons.** *Science* 2002, **295**:2258–61.

9. Schaeffer D, van Hoof A: **Different nuclease requirements for exosome-mediated degradation of normal and nonstop mRNAs.** *Proc Natl Acad Sci U S A* 2011, **108**:2366–71.

10. Klauer a A, van Hoof A: **Degradation of mRNAs that lack a stop codon: a decade of nonstop progress.** *Wiley Interdiscip Rev RNA* 2012, **3**:649–60.

11. Graille M, Séraphin B: **Surveillance pathways rescuing eukaryotic ribosomes lost in translation.** *Nat Rev Mol Cell Biol* 2012, **13**:727–35.

12. Shoemaker CJ, Eyler DE, Green R: **Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay.** *Science* 2010, **330**:369–72.

13. Doma MK, Parker R: **Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation.** *Nature* 2006, **440**:561–4.

14. Isken O, Maquat LE: **Quality control of eukaryotic mRNA: safeguarding cells from abnormal mRNA function.** *Genes Dev* 2007, **21**:1833–56.

15. Rebbapragada I, Lykke-Andersen J: **Execution of nonsense-mediated mRNA decay: what defines a substrate?** *Curr Opin Cell Biol* 2009, **21**:394–402.

16. Nicholson P, Mühlemann O: **Cutting the nonsense: the degradation of PTC-containing mRNAs.** *Biochem Soc Trans* 2010, **38**:1615–20.

17. LeBlanc JJ, Weil JE, Beemon K: **Posttranscriptional regulation of retroviral gene expression: primary RNA transcripts play three roles as pre-mRNA, mRNA, and genomic RNA.** *Wiley Interdiscip Rev RNA* 2013, **4**:567–80.

18. He F, Li X, Spatrick P, Casillo R, Dong S, Jacobson A: **Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5′ to 3′ mRNA decay pathways in yeast.** *Mol Cell* 2003, **12**:1439–52.

19. Mendell JT, Sharifi NA, Meyers JL, Martinez-Murillo F, Dietz HC: **Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise.** *Nat Genet* 2004, **36**:1073–8.

20. Kervestin S, Jacobson A: **NMD: a multifaceted response to premature translational termination.** *Nat Rev Mol Cell Biol* 2012, **13**:700–12.

21. Schweingruber C, Rufener SC, Zünd D, Yamashita A, Mühlemann O: **Nonsense-mediated mRNA decay - mechanisms of substrate mRNA recognition and degradation in mammalian cells.** *Biochim Biophys Acta* 2013, **1829**:612–23.

22. Barker GF, Beemon K: **Rous sarcoma virus RNA stability requires an open reading frame in the gag gene and sequences downstream of the gag-pol junction.** *Mol Cell Biol* 1994, **14**:1986–96.

23. Schoenberg DR, Maquat LE: **Regulation of cytoplasmic mRNA decay.** *Nat Rev Genet* 2012, **13**:246–59.

24. Shoemaker CJ, Green R: **Translation drives mRNA quality control.** *Nat Struct Mol Biol* 2012, **19**:594–601.

25. Franks TM, Singh G, Lykke-Andersen J: **Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense- mediated mRNA decay.** *Cell* 2010, **143**:938–50.

26. Kashima I, Yamashita A, Izumi N, Kataoka N, Morishita R, Hoshino S, Ohno M, Dreyfuss G, Ohno S: **Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay.** *Genes Dev* 2006, **20**:355–67.

27. Applequist SE, Selg M, Raman C, Jäck HM: **Cloning and characterization of HUPF1, a human homolog of the Saccharomyces cerevisiae nonsense mRNA-reducing UPF1 protein.** *Nucleic Acids Res* 1997, **25**:814–21.

28. Bhattacharya A, Czaplinski K, Trifillis P, He F, Jacobson A, Peltz SW: **Characterization of the biochemical properties of the human Upf1 gene product that is involved in nonsense-mediated mRNA decay.** *RNA* 2000, **6**:1226–35.

29. Metze S, Herzog VA, Ruepp M-D, Mühlemann O: **Comparison of EJC-enhanced and EJC-independent NMD in human cells reveals two partially redundant degradation pathways.** *RNA* 2013, **19**:1432–48.

30. Le Hir H, Izaurralde E, Maquat LE, Moore MJ: **The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions.** *EMBO J* 2000, **19**:6860–9.

31. Hir Â Le, Gat D, Izaurralde E, Moore MJ, Le Hir H, Gatfield D: **The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay**. *EMBO J* 2001, **20**:4987–4997.

32. He F, Brown AH, Jacobson A, Library M, Hopkins J: **Upf1p , Nmd2p , and Upf3p Are Interacting Components of the Yeast Nonsense-Mediated mRNA Decay Pathway**. 1997, **17**:1580–1594.

33. Withers JB, Beemon KL: **The structure and function of the Rous sarcoma virus RNA stability element.** *J Cell Biochem* 2011.

34. Weischenfeldt J, Waage J, Tian G, Zhao J, Damgaard I, Jakobsen JS, Kristiansen K, Krogh A, Wang J, Porse BT: **Mammalian tissues defective in nonsense-mediated mRNA decay display highly aberrant splicing patterns.** *Genome Biol* 2012, **13**:R35.

35. McIlwain DR, Pan Q, Reilly PT, Elia AJ, McCracken S, Wakeham AC, Itie-Youten A, Blencowe BJ, Mak TW: **Smg1 is required for embryogenesis and regulates diverse genes via alternative splicing coupled to nonsense-mediated mRNA decay.** *Proc Natl Acad Sci U S A* 2010, **107**:12186–91.

36. Roy B, Jacobson A: **The intimate relationships of mRNA decay and translation.** *Trends Genet* 2013:1–9.

37. Saulière J, Murigneux V, Wang Z, Marquenet E, Barbosa I, Le Tonquèze O, Audic Y, Paillard L, Roest Crollius H, Le Hir H: **CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex.** *Nat Struct Mol Biol* 2012, **19**:1124–31.

38. Singh G, Kucukural A, Cenik C, Leszyk JD, Shaffer S a, Weng Z, Moore MJ: **The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus.** *Cell* 2012, **151**:750–64.

39. Zhang Z, Krainer AR: **Involvement of SR proteins in mRNA surveillance.** *Mol Cell* 2004, **16**:597–607.

40. Balistreri G, Horvath P, Schweingruber C, Zünd D, McInerney G, Merits A, Mühlemann O, Azzalin C, Helenius A: **The Host Nonsense-Mediated mRNA Decay Pathway Restricts Mammalian RNA Virus Replication**. *Cell Host Microbe* 2014, **16**:403–411.

41. Muhlrad D, Parker R: **Aberrant mRNAs with extended 3´ UTRs are substrates for rapid degradation by mRNA surveillance.** *RNA* 1999, **5**:1299–307.

42. Amrani N, Ganesan R, Kervestin S, Mangus D a, Ghosh S, Jacobson A: **A faux 3′-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay.** *Nature* 2004, **432**:112–8.

43. Behm-Ansmant I, Gatfield D, Rehwinkel J, Hilgers V, Izaurralde E: **A conserved role for cytoplasmic poly(A)-binding protein 1 (PABPC1) in nonsense-mediated mRNA decay.** *EMBO J* 2007, **26**:1591–601.

44. Eberle AB, Stalder L, Mathys H, Orozco RZ, Mühlemann O: **Posttranscriptional gene regulation by spatial rearrangement of the 3′ untranslated region.** *PLoS Biol* 2008, **6**:e92.

45. Singh G, Rebbapragada I, Lykke-Andersen J: **A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay.** *PLoS Biol* 2008, **6**:e111.

46. Hogg JR, Goff SP: **Upf1 senses 3′UTR length to potentiate mRNA decay.** *Cell* 2010, **143**:379–89.

47. Kurosaki T, Maquat LE: **Rules that govern UPF1 binding to mRNA 3′ UTRs.** *Proc Natl Acad Sci U S A* 2013, **110**:3357–62.

48. Zünd D, Gruber AR, Zavolan M, Mühlemann O: **Translation-dependent displacement of UPF1 from coding sequences causes its enrichment in 3′ UTRs.** *Nat Struct Mol Biol* 2013, **20**:936–43.

49. Hurt J a, Robertson AD, Burge CB: **Global analyses of UPF1 binding and function reveal expanded scope of nonsense-mediated mRNA decay.** *Genome Res* 2013, **23**:1636–50.

50. Hogg JR: **This message was inspected by Upf1: 3′UTR length sensing in mRNA quality control**. *Cell Cycle* 2011, **10**:372–373.

51. Wiestner A, Tehrani M, Chiorazzi M, Wright G, Gibellini F, Nakayama K, Liu H, Rosenwald A, Muller-Hermelink HK, Ott G, Chan WC, Greiner TC, Weisenburger DD, Vose J, Armitage JO, Gascoyne RD, Connors JM, Campo E, Montserrat E, Bosch F, Smeland EB, Kvaloy S, Holte H, Delabie J, Fisher RI, Grogan TM, Miller TP, Wilson WH, Jaffe ES, Staudt LM: **Point mutations and genomic deletions in CCND1 create stable truncated cyclin D1 mRNAs that are associated with increased proliferation rate and shorter survival.** *Blood* 2007, **109**:4599–606.

52. Majoros WH, Ohler U: **Spatial preferences of microRNA targets in 3′ untranslated regions.** *BMC Genomics* 2007, **8**:152.

53. Mayr C, Bartel DP: **Widespread shortening of 3′UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells.** *Cell* 2009, **138**:673–84.

54. Ji Z, Tian B: **Reprogramming of 3′ untranslated regions of mRNAs by alternative polyadenylation in generation of pluripotent stem cells from different cell types.** *PLoS One* 2009, **4**:e8419.

55. Sandberg R, Neilson JR, Sarma A, Sharp P a, Burge CB: **Proliferating cells express mRNAs with shortened 3′ untranslated regions and fewer microRNA target sites.** *Science* 2008, **320**:1643–7.

56. Ozsolak F, Kapranov P, Foissac S, Kim SW, Fishilevich E, Monaghan a P, John B, Milos PM: **Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation.** *Cell* 2010, **143**:1018–29.

57. Tian B, Hu J, Zhang H, Lutz CS: **A large-scale analysis of mRNA polyadenylation of human and mouse genes.** *Nucleic Acids Res* 2005, **33**:201–12.

58. Mocquet V, Neusiedler J, Rende F, Cluet D, Robin J-P, Terme J-M, Duc Dodon M, Wittmann J, Morris C, Le Hir H, Ciminale V, Jalinot P: **The human T-lymphotropic virus type 1 tax protein inhibits nonsense-mediated mRNA decay by interacting with INT6/EIF3E and UPF1.** *J Virol* 2012, **86**:7530–43.

59. Nakano K, Ando T, Yamagishi M, Yokoyama K, Ishida T, Ohsugi T, Tanaka Y, Brighty DW, Watanabe T: **Viral interference with host mRNA surveillance, the nonsense-mediated mRNA decay (NMD) pathway, through a new function of HTLV-1 Rex: implications for retroviral replication.** *Microbes Infect* 2013, **15**:491–505.

60. Arrigo S, Beemon K: **Regulation of Rous sarcoma virus RNA splicing and stability.** *Mol Cell Biol* 1988, **8**:4858–67.

61. Barker GF, Beemon K: **Nonsense codons within the Rous sarcoma virus gag gene decrease the stability of unspliced viral RNA.** *Mol Cell Biol* 1991, **11**:2760–8.

62. Hayward WS: **Size and genetic content of viral RNAs in avian oncovirus-infected cells.** *J Virol* 1977, **24**:47–63.

63. Maquat LE, Li X: **Mammalian heat shock p70 and histone H4 transcripts, which derive from naturally intronless genes, are immune to nonsense-mediated decay.** *RNA* 2001, **7**:445–56.

64. LeBlanc JJ, Beemon KL: **Unspliced Rous sarcoma virus genomic RNAs are translated and subjected to nonsense-mediated mRNA decay before packaging.** *J Virol* 2004, **78**:5139–46.

65. Weil JE, Beemon KL: **A 3′ UTR sequence stabilizes termination codons in the unspliced RNA of Rous sarcoma virus.** *RNA* 2006, **12**:102–10.

66. Withers JB, Beemon KL: **Structural features in the Rous sarcoma virus RNA stability element are necessary for sensing the correct termination codon.** *Retrovirology* 2010, **7**:65.

67. Weil JE, Hadjithomas M, Beemon KL: **Structural characterization of the Rous sarcoma virus RNA stability element.** *J Virol* 2009, **83**:2119–29.

68. Toma KG, Rebbapragada I, Durand S, Lykke-Andersen J: **Identification of elements in human long 3′ UTRs that inhibit nonsense-mediated decay.** *RNA* 2015, **21**:887–97.

69. Cook KB, Kazan H, Zuberi K, Morris Q, Hughes TR: **RBPDB: a database of RNA-binding specificities.** *Nucleic Acids Res* 2011, **39**(Database issue):D301–8.

70. Withers JB: **Rous sarcoma virus and human telomerase reverse transcriptase mRNA expression**. *Ph.D. Dissertation,* Johns Hopkins University; 2011(December).

71. Ge Z, Quek BL, Beemon K, Hogg JR: **PTBP1 excludes UPF1 to protect long 3′UTRs from nonsense-mediated mRNA decay**. Submitted, 2015.

72. Romanelli MG, Diani E, Lievens PM-J: **New insights into functional roles of the polypyrimidine tract-binding protein.** *Int J Mol Sci* 2013, **14**:22906–32.

73. Zuker M: **Mfold web server for nucleic acid folding and hybridization prediction**. *Nucleic Acids Res* 2003, **31**:3406–3415.

74. Weil JE: **An RNA element that stabilizes translation termination in rous sarcoma virus**. *Ph.D. Dissertation,* Johns Hopkins University; 2007.

75. Ajamian L, Abrahamyan L, Milev M, Ivanov P V, Kulozik AE, Gehring NH, Mouland AJ: **Unexpected roles for UPF1 in HIV-1 RNA metabolism and translation.** *RNA* 2008, **14**:914–27.

76. Paige JS, Wu KY, Jaffrey SR: **RNA mimics of green fluorescent protein.** *Science* 2011, **333**:642–6.

77. Strack RL, Disney MD, Jaffrey SR: **A superfolding Spinach2 reveals the dynamic nature of trinucleotide repeat-containing RNA.** *Nat Methods* 2013, **10**:1219–24.

78. Filonov GS, Moon JD, Svensen N, Jaffrey SR: **Broccoli: rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution.** *J Am Chem Soc* 2014, **136**:16299–308.

79. Watts JM, Dang KK, Gorelick RJ, Leonard CW, Bess JW, Swanstrom R, Burch CL, Weeks KM: **Architecture and secondary structure of an entire HIV-1 RNA genome.** *Nature* 2009, **460**:711–6.

80. Pollom E, Dang KK, Potter EL, Gorelick RJ, Burch CL, Weeks KM, Swanstrom R: **Comparison of SIV and HIV-1 genomic RNA structures reveals impact of sequence evolution on conserved and non-conserved structural motifs.** *PLoS Pathog* 2013, **9**:e1003294.

81. Paca RE, Ogert RA, Hibbert CS, Izaurralde E, Beemon KL: **Rous sarcoma virus DR posttranscriptional elements use a novel RNA export pathway.** *J Virol* 2000, **74**:9507–14.

82. Méric C, Gouilloud E, Spahr PF, Mtric C: **Mutations in Rous sarcoma virus nucleocapsid protein p12 (NC): deletions of Cys-His boxes.** *J Virol* 1988, **62**:3328–33.

83. Cadwell RC, Joyce GF: **Mutagenic PCR.** *PCR Methods Appl* 1994, **3**:S136–40.

84. Mendell JT, Rhys CMJ, Dietz HC: **Separable Roles for rent1 / hUpf1 in Altered Splicing and Decay of Nonsense Transcripts**. 2002, **298**(October):419–422.

85. Amir-Ahmady B, Boutz PL, Markovtsov V, Phillips ML, Black DL: **Exon repression by polypyrimidine tract binding protein.** *RNA* 2005, **11**:699–716.

CURRICULUM VITAE

Bao Lin Quek

Born: 12 May 1983, in Singapore
Email: baolin.quek@gmail.com
LinkedIn: https://www.linkedin.com/in/baolinquek

4000 N Charles St Apt 1206
Baltimore, MD 21218
(608) 320-9320

EDUCATION

Ph.D. Candidate, Cellular, Molecular, Developmental Biology and Biophysics
2008 – present
Johns Hopkins University; Baltimore, MD — Cumulative GPA: 3.55 / 4.00
Research focus: Molecular Virology
Expected completion date: 30 June 2015
GRE: Quantitative: 800 / 800 (94%)
Verbal: 660 / 800 (94%)
Analytical writing: 5.5 / 6.0 (88%)

Bachelor of Science, Natural Sciences — **2002 – 2005**
University of Wisconsin – Madison; Madison, WI — Cumulative GPA: 3.40 / 4.00
Major: Genetics — Major GPA: 3.44 / 4.00
Major: Economics — Major GPA: 3.73 / 4.00

RESEARCH EXPERIENCE

Karen Beemon Laboratory — **2011 – present**
Department of Biology, Johns Hopkins University
Ph.D. Thesis: Characterization of retroviral RNA stability elements.
Investigated how retroviruses (HIV, MMLV, RSV) stabilize their viral genomes and evade host cell degradation mechanisms. Methods include genomic mutagenesis, RNA quantification and structural analyses, protein binding assays, and knockdown experiments.

Samer Hattar Laboratory — **2009 – 2011**
Department of Biology, Johns Hopkins University
Determination of the molecular components of intrinsically photosensitive retinal ganglion cell axon guidance. Analysis of sleep behavioral patterns in a mouse model for Schizophrenia.

Ph.D. program laboratory rotations — **2008 – 2009**
Department of Biology, Johns Hopkins University
Mentors: Rejji Kuruvilla, Xin Chen, Alex Bortvin, Chen-Ming Fan, Marnie Halpern

Julius Adler Laboratory **2004 – 2005**
Department of Biochemistry, University of Wisconsin – Madison
Behavioral screening and characterization of *D. Melanogaster* with neurological defects leading to transgressions in decision-making. Ramifications of aging on decision-making in *D. Melanogaster*

Francisco Pelegri Laboratory **2003 – 2005**
Department of Genetics, University of Wisconsin – Madison
Phenotypic characterization and mapping of a maternal-effect mutant gene that causes lysis during epiboly

PHARMACEUTICAL INDUSTRY EXPERIENCE

PPD, Inc, Bioanalytical Laboratory; Middleton, WI **2005 – 2008**
Associate scientist *2006 – 2008*
Assistant scientist *2005 – 2006*
Bioanalysis of clinical trial samples in a GLP-compliant setting to advance molecule development from preclinical studies through all phases of clinical trials. Preparation and analysis of pharmaceutical compounds in biological matrices via LC/MS/MS, HPLC, and spectrophotometry. Responsibilities include report generation, project organization, method validation, instrument error troubleshooting, review of study and quality control data, and knowledge of OSHA and FDA GLP and cGMP regulations.

PUBLICATIONS

Ge Z, **Quek BL**, Beemon KL, Hogg JR. (2015, submitted) PTBP1 excludes UPF1 to protect long 3' UTRs from nonsense-mediated mRNA decay.

Quek BL, Beemon K. (2014) **Retroviral strategy to stabilize viral RNA**. *Curr. Opin. Microbiol. 18*, 78-82.

CONFERENCE PRESENTATIONS AND ATTENDANCE

18th Annual HIV Drug Resist. Prog. Think Tank Meeting; Frederick, MD 2015
Oral presentation: "Characterization of retroviral RNA stability elements"

National Cancer Institute Symposium; Bethesda, MD 2015
RNA Biology
Poster: "Characterization of retroviral RNA stability elements" Quek BL, Beemon K.

Gordon Research Conference; Newport, RI 2014
Post-Transcriptional Gene Regulation – From Mechanism Stems Therapy
Poster: "Characterization of retroviral RNA stability elements" Quek BL, Beemon K.

Gordon-Merck Research Seminar; Newport, RI 2014
Post-Transcriptional Gene Regulation – Controlling Gene Expression Through mRNA Regulation
Selected oral presentation: "Characterization of retroviral RNA stability elements"

17th Annual HIV Drug Resist. Prog. Think Tank Meeting; Frederick, MD 2014

Workshop on Post-Transcriptional Reg. of Virus Expression; Syria, VA 2013
Selected oral presentation: "Characterization of retroviral RNA stability elements"

Seventeenth Annual Meeting of the RNA Society 2012
Poster: "The Rous Sarcoma Virus RNA stability Element acts as an insulator to prevent recognition of unspliced retroviral RNA by host cell decay machinery" Withers JB, Quek BL, Ingolia NT, Beemon KL.

HIV Drug Resistance Program Conference; Frederick, MD 2012
Trafficking of Viral Macromolecules

HONORS / AWARDS

Dupont Award for Excellence in Teaching **2010**
Developmental Biology Laboratory, Johns Hopkins University.

Dean's List **2003 – 2005**
College of Agriculture and Life Sciences, University of Wisconsin – Madison.

TEACHING

Teaching Assistant, Virology course, JHU 2015, 2012

Laboratory Teaching Assistant, Developmental Biol. course, JHU 2010

Laboratory Teaching Assistant, Genetics course, JHU 2009

PROFESSIONAL AFFILIATIONS

RNA Society 2012

American Chemical Society 2006 – 2008

SELECTED LEADERSHIP EXPERIENCE

Johns Hopkins University Ballroom Dance Club; JHU 2008 – present
Founder 2008 – present
Executive council member 2008 – 2012
Instructor 2009 – 2013
Founded a new student organization. Successfully applied for and awarded multiple grants. Organized recruitment and fundraising drives, performances, and social dances. Liaised with other collegiate teams, as well as the national governing body for dancesport, to coordinate team participation in regional and national competitions. Excellent track record of soliciting donations, advertising, and community outreach. Trained hundreds of beginners and seasoned competitors.

Singapore Students Association; Univ. of Wisc – Madison 2002 – 2005
Treasurer 2003 – 2004
Successfully applied for and awarded operations and travel grants; incentivized and achieved a five-fold increase in seminar attendance rates.

OTHER SKILLS

Computer
Microsoft Windows and Office Suite
Macintosh Operating Systems and iWork Suite

Languages
Native proficiency in English
Native proficiency in Mandarin Chinese

Licenses and Certifications
Class C Driver's License
PADI Advanced Open Water Diver

Business
Founder and owner of Lin Quek Ballroom Dance Instruction and Choreography